Dynamic structure of the full-length scaffolding protein NHERF1 influences signaling complex assembly

Shibani Bhattacharya[1*], Christopher B. Stanley[2], William T. Heller[2], Peter A. Friedman[3*], and Zimei Bu[4*]

1. New York Structural Biology Center, 89 Convent Avenue, New York, NY, 10027, USA.
2. Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, TN
3. Department of Pharmacology and Chemical Biology, University of Pittsburgh, 200 Lothrop Street, Pittsburgh, PA 15261
4. Department of Chemistry and Biochemistry, City College of New York, CUNY, 160 Convent Avenue, New York, NY 10031

*Address correspondence to: sbhattacharya@nysbc.org, paf10@pitt.edu, or zbu@ccny.cuny.edu

Keywords: scaffolding protein, PDZ domain, G protein-coupled receptor kinase 6A (GRK6A), intrinsically disordered protein, NHERF1, NMR, small-angle neutron scattering (SANS), SLC9A3 regulator 1 (SLC9A3R1), intracellular transport

ABSTRACT

The Na^+/H^+ exchange regulatory cofactor 1 (NHERF1) protein modulates the assembly and intracellular trafficking of several transmembrane G-protein coupled receptors (GPCRs) and ion transport proteins with the membrane-cytoskeleton adapter protein Ezrin. Here, we applied solution NMR and small-angle neutron scattering (SANS) to structurally characterize full-length NHERF1 and disease-associated variants that are implicated in impaired phosphate homeostasis. Using NMR, we mapped the modular architecture of NHERF1, which is composed of two structurally-independent PDZ domains that are connected by a flexible, disordered linker. We observed that the ultralong and disordered C-terminal tail of NHERF1 has a Type 1 PDZ binding motif that interacts weakly with the proximal, second PDZ domain to form a dynamically autoinhibited structure. Using ensemble-optimized analysis of SANS data, we extracted the molecular size distribution of structures from the extensive conformational space sampled by the flexible chain. Our results revealed that NHERF1 is a diffuse ensemble of variable PDZ domain configurations and a disordered C-terminal tail. The joint NMR/SANS data analyses of three disease variants (L110V, R153Q and E225K) revealed significant differences in the local PDZ domain structures and in the global conformations compared to the wild-type protein. Further, we show that the substitutions affect the affinity and kinetics of NHERF1 binding to Ezrin and to a C-terminal peptide from GPCR kinase 6A (GRK6A). These findings provide important insight into the modulation of the intrinsic flexibility of NHERF1 by disease-associated point mutations that alters the dynamic assembly of signaling complexes.

Signal transduction in the biological milieu is a dynamic exchange of protein-

protein interactions (1,2), coordinated by a diverse family of scaffolding proteins (2,3,4). Structural modularity of the scaffolds is fine-tuned for linking and transporting the different binding partners to the supramolecular signaling complexes. Well known scaffold modules include SH3, SH2, PTB, WW and the more abundant PDZ domains, which are ubiquitous for recruiting diverse protein targets (3,4). As mediators of protein-protein interactions, PDZ domains have been shown to possess remarkable specificity for C-terminal binding motifs as well as internal peptide sequences of membrane proteins (3). PDZ domains have emerged as key organizers of protein complexes at the plasma membrane, cytoplasmic tails of membrane proteins, and ion channels to promote their transport and localization to the cell surface for signaling (5).

Full-length NHERF1 includes tandem PDZ domains, PDZ1 and PDZ2, and a C-terminal Ezrin binding domain (EBD) juxtaposed with a PDZ-binding Type 1 motif (**Figure 1**). The EBD is a specific target of the FERM domain from Ezrin (ezFERM), which disrupts the autoinhibition between the PDZ2 domain and the C-terminal tail in NHERF1 (6). Binding of Ezrin with NHERF1 allosterically increases the PDZ domain affinity for ligands, such as the transmembrane receptors and ion channels (7,8). Examples of transmembrane protein complexes that are regulated by the NHERF1-Ezrin linkers include the transmembrane sodium–hydrogen exchanger 3 (NHE3) complexes, CFTR (9), the parathyroid hormone receptor (10), and the CHIP-based protein quality control machinery (11).

Previously, we employed small-angle X-ray and neutron scattering, SAXS and SANS, respectively, to investigate the global conformational switch triggered in full-length NHERF1 by association with ezFERM (7). The study revealed a mesoscopic movement of the PDZ domains upon switching between the autoinhibited 'closed' structure (R_g=41 Å, D_{max}=140 Å) and the target-accessible 'open' conformation (R_g=51 Å, D_{max}=175 Å) (7,12,13). However, the earlier results lacked the resolution required to correlate the ensemble dependent dynamics of the modular structure with the extent of order in the ultra-long spacers. The interdomain linkers constitute nearly 30% of the NHERF1 sequence and are remarkably well conserved (>90%) across different species (10) but possesses low structural propensity (**Figure 1**).

Traditional NMR has been highly successful in characterizing flexible scaffolding proteins with ordered domains randomized by unstructured linkers (14,15). The high-resolution solution structures of the isolated PDZ1 and PDZ2 domains from NHERF1 with novel helical extensions that stabilize the structure and allosterically modulate peptide affinity (16,17). Here, multi-dimensional solution NMR combined with SANS experiments were used to study the complete structural and dynamic characteristics of the intact protein. Our results reveal full-length NHERF1 cannot be represented by a single conformation. Instead, it is a dynamic ensemble of variable PDZ domain configurations linked by highly flexible regions.

Naturally occurring mutations in human *NHERF1* gene have been implicated in impaired renal phosphate transport with increased risk of nephrolithiasis, bone demineralization (18), and cancer (19). Gene sequencing of afflicted patients identified at least three loss-of-function mutations (L110V, R153Q and E225K) in NHERF1 (18). In the isolated PDZ1 (L110V) and PDZ2 (R153Q) domains, we have shown the mutations modulate the target affinity by lowering the thermodynamic stability of the protein (16,17). In the present study, we investigated the mutations in the context of full-length

NHERF1, to elucidate their simultaneous impact on the intramolecular autoinhibition of the PDZ2 domain and target affinity. Using NMR, we mapped the local perturbation in each PDZ domain and, combined with ensemble analysis of SANS data, demonstrated the effect of the mutations on the global structure and function.

Ezrin is indispensable for releasing autoinhibition, but phosphorylation plays an equally critical role in modulating the cellular activity of NHERF1 through its ability to oligomerize (20) and associate with various binding partners (10,21,22). In humans, NHERF1 is reversibly phosphorylated at Ser290 by a multiprotein kinase cascade involving GRK6A, which in turn regulates PTH-sensitive, NPT2A-mediated phosphate uptake (23,24). Although GRK6A has much higher affinity for the N-terminal PDZ1 domain, it is through interactions with the partially masked PDZ2 domain that NHERF1 is phosphorylated in the C-terminal hinge region (20). We have further evaluated the extent to which the structural differences affect disease mutations to interact with Ezrin and G protein-coupled receptor Kinase A (GRK6A). We used SPR to measure the binding affinity and the kinetics of the association of each of the three mutants with ezFERM and a C-terminal peptide from GRK6A. Collectively, the results provide a structural and dynamic framework for understanding the internal mechanics of wild-type and NHERF1 variants in key functional states.

RESULTS

NMR Structural Analysis of full-length NHERF1

Using structural information provided by NMR chemical shift analysis, we confirmed the boundaries of the structured domains, nature of interdomain contacts, and the structural propensities of the linker regions in full-length NHERF1. The initial characterization focused on the autoinhibited and 'open' state triggered by association of NHERF1(wt) with the FERM domain of Ezrin (ezFERM). Despite the relatively high molecular mass of NHERF1 (~39 kDa) and its complex with ezFERM (~79 kDa), the high quality of the 2D N^{15}-TROSY spectra (**Figure 2A**) enabled us to successfully assign ~76% of the backbone resonances of the 358-residue construct, using $^{2}H/^{13}C/^{15}N$-labeled samples and TROSY-based triple resonance experiments.

In NHERF1(wt), the 'on-off' association of PDZ2 with the carboxy terminal residues results in extensive broadening of backbone amide resonances in both PDZ2 (α2, β2-β3), and EB (residues 337-358) domains due to coalescence of resonances on intermediate timescales. In the ezFERM-bound state, the conformational exchange in the PDZ2 domain is quenched, and some of the largest chemical shift changes mapped to residues in the ligand binding site (**Figure 2B**). Unexpectedly, we also observed line-broadening in the β2 strand (residues 25-31) in the PDZ1 ligand binding site. Like the PDZ2 domain, we recovered the intensity of the lost signals in PDZ1 upon ezFERM binding. Collectively these changes are consistent with transient autoinhibitory interactions involving the PDZ2 domain, and to a lesser extent, the PDZ1 domain.

The secondary structure predictions obtained from the analysis of the backbone chemical shifts (Cα, Cβ, C', N and H^{N}) of NHERF1 in the autoinhibited state using TALOS+ (25) are displayed in **Figure 2C**. The secondary structure profile is identical to the αβ fold of the isolated PDZ domains (16,17), which is preserved upon ezFERM-binding in the 'open' state (**Figure S1**). The random coil index (RCI) from TALOS+ analysis indicates

that the linker regions, connecting PDZ1-PDZ2 domains (residues 112-146) and between PDZ2-EB domains (residues 252-324), have disordered backbones (**Figure S1**). To confirm the RCI predictions, we compared the experimental chemical shifts against the random coil values incorporated in the POTENCI database used for identifying intrinsically disordered proteins (IDPs) (26). The excellent agreement between the predicted random coil shifts of the linker residues and the corresponding experimental values confirmed that these regions are intrinsically disordered and completely flexible (**Figure 3**). The unusually long (residues 262-324) and disordered C-terminal tail lends credence to the possibility that the EBD could interact with both PDZ domains, albeit with different affinities.

Dynamic intramolecular autoregulation in NHERF1.

In the isolated PDZ1 domain, the residues in the ligand-binding site can be observed in the 2D N^{15}-HSQC spectrum (17). Hence, the logical origin of the observed loss-of-signal in PDZ1 from full-length NHERF1 is either a *weak* intra- or inter-molecular interaction with an internal binding motif or the carboxy terminal sequence (-**FSNL**). Because we recover the lost signal intensity upon the EB domain binding to ezFERM, we can disqualify purported interactions involving an internal segment in the linker regions. Instead, the most probable cause is a weak *intramolecular* interaction with the C-terminal Type 1 motif (-**S**x**L**), which is effectively independent of the protein concentration (>40 μM). The C-terminal motif without a positively charged side-chain at the penultimate position (x=Arg^{+}) is not optimized for high-affinity binding to the PDZ1 domain (17,27). Nevertheless, we expect the weak millimolar affinity could increase by the local concentration effect.

Thus, in solution, monomeric NHERF1 can adopt at least two inactive 'closed' conformations and a fully 'open' state when bound to ezFERM. NHERF1 has been shown to dimerize at concentrations greater than 140 μM that are physiologically (<10 μM) irrelevant (13,22,28). To estimate the contribution of the different structural states of NHERF1 to any experimental measurement, we applied a quantitative relationship between intra- and intermolecular binding constants developed originally for polymer chains of variable length in systems of similar complexity (29) (30) (31):

$$K_i' = C_{eff} \times K_a^i \qquad (1)$$

$$f_{bound}^i = \frac{K_i'}{(1+\sum K_i')}, f_{free} = \frac{1}{(1+\sum K_i')}$$

where K_a^i is the i^{th} bimolecular association constant, C_{eff} is the local concentration calculated from the end-to-end distance probability distribution (**Figure S2**), and K_i' is the effective association constant. Based on the bimolecular affinity (K_d ~ 93 μM) between the isolated PDZ2 domain and a C-terminal peptide (32), we expect at least ~77% monomer at the protein concentrations (~140 μM) required for the NMR studies. Using the worm-chain model, the probability of the end-to-end distance of the 91-residue linker between PDZ2 and EB domain maximizes at ~37 Å. In this distance range, free NHERF1 is a mixed population of 'closed' (76%) and 'open' (24%) conformations as predicted by the intramolecular association constant K' and effective local concentration, C_{eff} (**Equation 1 and Figure S2**). Assuming a slightly longer end-to-end distance of the linker (~60 Å) between the PDZ1-EB domains based on model structures, the effective local concentration is lower, resulting in a much

smaller fraction of 'closed' PDZ1 conformer (<5%).

The significant population of the 'open' conformer of NHERF1 (~24%) has important biological ramifications for access to the target binding site at both PDZ domains when Ezrin is dormant. Ezrin in the inactive state is autoinhibited, masking the NHERF1 binding site which is only exposed by a conformational change triggered by phosphorylation and PIP_2 binding (33,34). To illustrate the effect of the interconverting NHERF1 populations on binding site access, we used a 20-residue C-terminal peptide derived from GRK6A (**Figure 4**) to simultaneously monitor the chemical shift perturbation in each PDZ domain at increasing peptide concentrations in the absence of ezFERM. The GRK6A peptide has much higher binding affinity for the isolated PDZ1 compared to the PDZ2 domain from NHERF1 (23). Accordingly, the N-terminal PDZ1 domain was readily saturated at 1:1 protein-to-peptide ratio. In the PDZ2 domain, binding was inhibited but not abrogated, supporting limited access permitted by the dynamic C-terminal interactions (**Figure 4B**). Therefore, an important outcome of partially 'open' states is different ligands have some access to both PDZ domains instead of relying on the complete release of autoinhibition by ezFERM binding. The dynamic mechanism offers a plausible explanation for why the PDZ2 domain is essential for phosphorylating Ser290 in the C-terminal domain of wild-type NHERF1 (23) (24).

Effects of disease mutations on local structure of NHERF1

To map the structural differences between NHERF1(wt) and the disease-associated variants (L110V, R153Q and E225K), we obtained sequence-specific resonance assignments of each protein. The corresponding amide ($^{1}H/^{15}N$) chemical shift differences (CSD) compared to the wild-type protein are plotted in **Figure 5**. In the histogram plots, we observe distinct clusters of residues with significant CSD in the PDZ1 (**Figure 5A-5B**) and PDZ2 domains (**Figure 5D-5E**, **5G-5H**) that correlate directly with the location of the mutations in the sequence. The structural integrity of the PDZ domains was confirmed by analyzing the backbone resonances in TALOS+, which did not disclose any unexpected loss of secondary structure (**Figure S3**). The canonical fold, consisting of a six-stranded β-sheet and four α-helices, is preserved in all three mutants (**Figure S3**).

In the L110V mutant, the structural perturbation is localized mainly in the C-terminal helical extension of the PDZ1 domain and the N-terminus of the linker (**Figure 5A-5B**). Based on the PDZ2 domain structure (16,35), R153Q mutation is expected to disrupt favorable electrostatic interactions between the side-chain of R153 $N^{\eta+}$ (β1) and negatively charged COO- groups from D197 (β4) and D232 (β6). As a consequence any rearrangement in the side-chain packing results in chemical shift changes across strands of the β-sheet structure (**Figure 5E**). When compared to the wild-type protein, the E225K mutation results in very small amide chemical shift changes along the backbone (<0.06 ppm) suggesting limited structural perturbation (**Figure 5G**). However, the location of the charge reversal mutation near the CB loop (facing out) is likely to have a direct effect on the CB loop interactions with the carboxylate group (**Figure 5H**).

Based on the extent of structural changes observed within the PDZ domains, we were interested in elucidating how these effects propagate to the binding site and mediate intramolecular interactions. Specifically, the two mutations in the PDZ2 domain are likely to influence the conformational switch between the 'closed' and 'open' states. The most suitable NMR probes to follow the changes in

the intramolecular switch would be the chemical shifts of residues in the EB domain or those in the carboxylate binding loop (-^{163}GYGF^{166}-). Because neither mutation had a detectable effect on the intermediate exchange phenomenon that bleaches out most of the NMR resonances from the residues in the binding site, we focused instead on Gly23 from PDZ1 and Gly163 from the PDZ2 domain (**Figure 1**). The two Glys represent the sole NMR reporters within the binding loop that offers critical structural and dynamic information on interactions with ligands.

The effect of the binding loop configurations on the two NMR probes (Gly23/Gly163) is shown in the overlay of the wild-type (close/open), ezFERM bound (open) and the mutant NHERF1 proteins (**Figure 5C/5F/5I**). The release of wild-type PDZ2 autoinhibition triggered by ezFERM binding is clearly captured by a shifted Gly163 amide crosspeak (blue contours) in the NHERF1(wt) 2D N^{15}-TROSY spectrum. In contrast, Gly23 in PDZ1 domain remains unaffected by the C-terminal conformational changes. Compared to the wild-type protein (black contours), none of the mutations display a significant chemical shift change at Gly163 (red contours), which would be consistent with the fully 'open' state analogous to the ezFERM bound protein. Instead we observe differential $^{1}H^{N}$ line-broadening that reflects a dynamic shift in the exchange phenomenon in the CB loop in both the PDZ domains. The elevated $^{1}H^{N}$ linewidth of Gly163 in E225K (~63 Hz) compared to L110V/R153Q mutants (~45.5±2.5 Hz) and the wild-type protein (~44 Hz) suggests this mutation near the CB loop witnesses an unmediated effect of the C-terminal interactions (**Figure 5H**). In the fully open ezFERM-bound state of NHERF1(wt), the much sharper $^{1}H^{N}$ linewidths (~21 Hz) indicate the lack of conformational exchange in the unbound CB loop from the PDZ2 domain. The overall trend at Gly23 suggests increased $^{1}H^{N}$ line broadening in the mutants (~32±1 Hz) compared to the autoinhibited (~22 Hz) and 'open' state (~26 Hz). It is very likely that the altered dynamics at Gly23 in the PDZ1 domain indicates a subtle shift in the population between the different autoinhibited configurations in the variants. It is reasonable to assume that the exchange rate itself is unchanged because the mutations are structurally remote and cannot directly mediate the interactions at the PDZ1 CB loop. In summary, the NMR results evince the mutations fail to deactivate the conformational switch wholly and the CB loop in the PDZ2 domain is interconverting between a variable fraction of C-terminal 'bound' and 'free' states respectively (16,32).

To further explore ligand access to both PDZ domains in the mutants, we also compared the CB loop Gly23/Gly163 chemical shifts in the presence of a peptide from GRK6A (**Figure S4**). As evinced by the conspicuous chemical shift changes in Gly23, all three mutants bind the peptide in the PDZ1 domain like the wild-type protein at 1:1 protein-to-peptide ratio (**Figure 4B**). In the PDZ2 domain, the broad Gly163 peak suggests it is not saturated at even 1:4 ratio (**Figure S4D-S4F**). Thus, we have established that the mutations have a subdued impact on the structural mechanism of intramolecular autoinhibition. Owing to adverse NMR lineshapes, any shift in the 'open' and 'closed' populations could not be reliably quantified. Therefore, we decided to pursue a different approach using SANS data to study the conformational properties of wild-type and mutant NHERF1 protein as a dynamic ensemble in solution.

Global conformational changes in wild-type and mutant NHERF1 from SANS.

The SANS experiments on the wild-type and variants of ^{d}NHERF1 (deuterated) were performed at low protein concentrations (~25 μM), where it is predominantly

monomeric (13) (**Table 1** and **Figure 6A-6B**). The corresponding P(r) distribution function are similar reinforcing the structural similarities between the wild-type protein and the mutants (**Figure 6C**).

To tease out the conformational differences between the NHERF1 variants, we performed an ensemble analysis of the SANS data using the program EOM (36,37) by generating an ensemble pool consisting of the major 'closed' and 'open' conformations. Several programs are available for fitting solution scattering data to large ensembles of multi-domain proteins (36,38,39). However, an inherent limitation of the existing software is the inability to sample the configurational space of internal loops with restrained termini. Therefore, to sample the 'closed' conformation of NHERF1 we used rigid body Monte Carlo (MC) simulation of PDZ domain motion by allowing the linkers to break during the random spatial search in Xplor-NIH (14). The linkers were rejoined, and the covalent geometry fixed before performing simulated annealing in torsion angle space (TAD). A more detailed description of the protocol used to generate the ensemble pool of the 'closed' and 'open' conformations of NHERF1 is provided in the Experimental procedures section. As seen in **Figure 6D-6E**, the width of the starting ensemble occupies a significant volume of the conformational space using a combination of MC and TAD in Xplor-NIH. The results of the ensemble-averaged analysis of the SANS data in the program GAJOE from EOM (36,37) are summarized in **Figure 6** and **Figure S5**. The quality of the ensemble-averaged fit of the calculated scattering intensities to the experimental SANS data ($\chi^2 < 0.8$) was quite good (**Table 1**).

The overall molecular size distribution profile of the wild-type protein that fits the SANS data is consistent with the SEC-SAXS data (**Table 1**) of monomeric NHERF1 reported in our previous studies (40). The close agreement between the R_g values confirms deuteration does not affect the conformation of NHERF1. From the overlays in **Figure 6D-6E**, the difference in the ensemble properties of the three mutants compared to wild-type protein is quite significant. In **Figures 6F-6K** representative structures from the 'optimized ensemble' from GAJOE (36) are displayed along with dimensions and contributions from pool fractions. R_{flex} is the metric used to describe the flexibility of the *selected* ($R_{flex,selected}$) ensemble compared to that of the *random* starting pool. It can vary from 100% (flexible) to a perfectly rigid system (0%). Likewise, the ratio of the variance of the selected/starting ensemble distribution, $R_{sigma}<1$ indicates flexibility. As summarized in **Table 1**, the relatively high $R_{flex,selected}$=83% and R_{sigma}=0.90 values, reflect the broad distribution of a highly flexible system sampling simultaneously partially 'open' and 'closed' structures in solution. There is a significant fluctuation in the distance between the two PDZ domains, which is proportional to the presence of compact (R_g~47 Å) and more extended chains (R_g~75 Å). Owing to the length of the ultralong and flexible linkers, the pool fraction was uncorrelated with the D_{max} values or state of the intramolecular association (**Figure 6F-6K**). Likewise, the ensemble populations bear limited correlation with the theoretically predicted ratio of 'open' and 'closed' states. Thus, the apparent randomness of the rotational and translational motion of the linked PDZ domains suggests purely Brownian dynamics is at play in the absence of persistent interactions to restrict the domain motion (14,33)(41). This conclusion is well supported by the NMR data where the two PDZ domains are structurally independent and flexibly tethered by intrinsically disordered linkers.

By NMR we demonstrated the overall β-sheet fold in the mutated PDZ domains is preserved and thus we can assume the backbone trace is also similar to the wild-type protein. Because our goal was to evaluate general trends in the R_g and D_{max} profiles for

the mutants, we fitted the experimental SANS data using the same random ensemble pool as NHERF1(wt) in GAJOE. In the overlay of the R_g and D_{max} distribution profiles, the L110V mutation in the PDZ1 domain clearly behaves differently from the other two mutations located in the PDZ2 domain (**Figure 6, Table 1**). The higher D_{max} values in L110V mutant, can be rationalized by structural changes at the junction between the extended PDZ1 structure and the disordered linker (**Figure 5A**). The mutation in the dynamic C-terminal 3^{10}-Helix (17) could modulate the elasticity of the linker changing the effective distance between the PDZ domains.

Compared to wild-type protein, the mutated PDZ2 domains (R153Q/E225K) favor a narrow matching R_g and D_{max} distribution profiles. This trend is also reflected in the reduced R_{flex} (~71%) and R_{sigma} (~0.61) values (**Table 1**). The major fraction (~80%) of the structures in the unbiased ensemble selected by GAJOE to fit the experimental data appear to favor randomly oriented C-terminal tails. In the E225K mutant, the average separation between the PDZ domains in the dominant conformer shrinks dramatically (R_g~45 Å) from NHERF1(wt) (R_g~75 Å) (**Figure 6K**). There is some evidence to suggest the unstructured C-terminal tail has a propensity to collapse when released from interactions with the PDZ2 domain (40). Thus, it is possible the compact ensemble reflects the population is shifting either in favor of partially 'open' states with collapsed chains or alternate weak interactions at the PDZ1 domain as evidenced by the NMR analysis.

Binding of ezFERM to NHERF1 mutants

To understood the impact of the disease mutations on the biochemical functions of full-length NHERF1, we have used SPR to analyze the binding affinity and kinetics with ezFERM and the 22 residue peptide of the C-terminal tail of GRK6A that possesses a type 1 PDZ-binding motif -PTRL (23).

In **Figure 7A**, the equilibrium binding curves of NHERF1 or mutants to ezFERM can best be fit by a 1:1 binding model with low χ^2 values (<**5.0**). Overall, NHERF1(wt) and the disease mutants have comparable equilibrium dissociation constants (**Table S1**) in the 13-43 nM range, with some variation amongst the mutants. We have shown previously that the SPR sensorgrams with Ezrin (**Figure S6**) are better fit by a two-step kinetic model (40) involving a fast and slow process (**Equation 2**), instead of a single 1:1 state or the bivalent binding model.

$$A + B \underset{k_{d1}}{\overset{k_{a1}}{\rightleftharpoons}} AB^* \underset{k_{d2}}{\overset{k_{a2}}{\rightleftharpoons}} AB \qquad (2)$$

In equation 2, A refers to ezFERM and B to NHERF1 or a mutant, k_{a1}/k_{d1} to the association/dissociation rate of the fast binding event, and k_{a2}/k_{d2} the corresponding rate constants for the slow binding event.

In NHERF1(wt), the fast binding mode has an association rate constant k_{a1} =1.58x10^7 $M^{-1}s^{-1}$, similar to values reported earlier (42). In the L110V mutant, the association rate k_{a1} is higher than wild-type and other mutants by an order of magnitude (**Figure 7B**). The fast association is correlated with the increased separation between the PDZ and EB domains (**Figure 6E**). Thus, making the EB domain associated with PDZ2 domain more accessible to ezFERM. The off-rate, k_{d1} shows small variation but is not greater than an order of magnitude across the different proteins.

The slower association rate k_{a2} is about 2-3 orders of magnitude smaller than k_{a1} for the wild-type protein with a monotonic decrease in the mutants (**Figure 7C**). We attribute the slower kinetic process to conformational

changes in the C-terminal tail of NHERF1 upon docking with ezFERM accompanied by weak binding interaction (16,40,43). This two-step kinetic model is well supported by the published X-ray structure of the 38-residue EB domain (a.a. 321–358) from NHERF1 bound to the moesin FERM domain (43). The X-ray structure revealed a high affinity C-terminal α-helix (residues 348-358) bound to the F3 subdomain of FERM linked to a second N-terminal helix (residues 325-334) which interacts weakly with the F2 subdomain of FERM (44). The monotonic decrease in the k_{a2} values in the mutants could very well be due to increased disorder in the EB domain because of a shift in the population between the 'closed' and 'open' states (16).

Binding of GRK6A peptide to NHERF1 mutants

The equilibrium binding curves of NHERF1(wt) or mutants to GRK6A peptide were best fit with a bivalent model. This model also makes the most physical sense, given the tandem PDZ binding sites. The bivalent kinetic model that fits SPR sensorgrams can be described as follows:

$$2A + B_1B_2 \underset{k_{d1}}{\overset{k_{a1}}{\rightleftharpoons}} AB_1B_2 + A \underset{k_{d2}}{\overset{k_{a2}}{\rightleftharpoons}} A_2B_1B_2 \qquad (3)$$

where A is the GRK6A peptide, B1 and B2 refer to PDZ1 and PDZ2 domains in NHERF1, k_{a1}/k_{a2} are bivalent association rate constants, and k_{d1}/k_{d2} are the bivalent dissociation rate constants.

In NHERF1(wt) we obtain effective dissociation constants K_{d1} ~ 160 nM and K_{d2} ~ 4000 nM by fitting the data with the bivalent binding model. As shown by NMR, the difference in affinity reflects the competition of the GRK6A peptide to displace the C-terminal residues from the PDZ2 binding site. The equilibrium K_{d1} values show very little variation amongst the mutants (**Figure 8A, Table S2**), with L110V and R153Q mutants possessing slightly higher affinity than E225K and NHERF1(wt). The K_{d2} values are nearly constant. Overall the equilibrium dissociation constants are consistent with what we expect from the peptide sequence analysis. The intrinsic affinity of the GRK6A peptide for the wild-type PDZ1 domain should be higher than the PDZ2 domain (-**PTRL**) owing to the positively charged Arg^+ located at the penultimate site (-1) [15].

The kinetic measurements confirmed both NHERF1(wt) PDZ domains bind the GRK6A peptide but with significantly different association rate constants. The fast association (k_{a1} ~$1.2x10^7$ $M^{-1}s^{-1}$) value for NHERF1(wt) is typical of a diffusion-limited process (17,45,46). Increased accessibility of the PDZ1 ligand binding site in NHERF1 irrespective of the state of autoinhibition, results in at least 3 orders of higher association rate than the PDZ2 domain (**Figure 8B, 8D**). As shown in **Figure 8B** all three mutations retard the association rate constant by an order of magnitude compared to NHERF1(wt) at the PDZ1 domain. The k_{d1} of R153Q and E225K are comparable to the wild-type protein, but that of L110V is apparently slower than NHERF1(wt). In the PDZ2 domain, the R153Q and E225K mutations results in slower k_{a2} and k_{d2} values which deviate from NHERF1(wt). The slower kinetics primarily reflect autoinhibition and structural changes within the PDZ2 binding site that impede the formation of the complex with GRK6A and interfere with phosphorylation.

DISCUSSION

The modular structure of NHERF1 is highly advantageous for its scaffolding function in assembling and trafficking of cellular complexes. The flexibility of the modular architecture offers a powerful

mechanism to permutate a single chain configuration into an ensemble of structures with minimum entropic barrier (47,48). As a prototypical scaffold we established that NHERF1 exists as a continuum of structures that can coordinate the spatial and temporal location of the PDZ1 and PDZ2 domains with respect to each other. The function of the intact protein can no longer be described by a binary switch shuttling exclusively between a 'closed' and 'open' conformation. Instead, it can be viewed in solution as a diffuse ensemble of different PDZ and EB domain configurations. This dynamic description contrasts with structurally similar multi-PDZ domain proteins (49) that have limited mobility, where short linkers result in correlated motion of the domains. In NHERF1, we observe a range of interdomain distances (40-80 Å) and random alignment within the conformational space populated by the ensemble that would support nearly independent movement of the PDZ domains despite being linked.

The length and number of linkers is a common evolutionary strategy used to modify the conformational space accessible to modular proteins (50). By NMR we confirmed nearly 30% of the NHERF1 chain is disordered, which agrees with the amino acid compositional bias in the linker and C-terminal region. For example, the linker region between the PDZ domains is dominated by disorder promoting residues like Pro (20%) and charge rich Glu (20%) (51,52). While the C-terminal tail is relatively rich in Ser (20%) with tunable flexibility by phosphorylation (53-55). Since NHERF1 is primarily involved in cross-linking Ezrin bound to the actin cytoskeleton and plasma membrane proteins, the ability to sample an unusually large conformational space has an innate advantage for capturing targets and bridging the different components of the signaling pathways.

The length and overall flexibility of the autoinhibitory C-terminal tail carrying the PDZ binding motif is vital for modulating the strength of the intramolecular interaction with the PDZ2 domain. Based on the structural models used to fit the SANS data, the effective chain length corresponding to 91 residues (~40 Å) appears to be within a range suitable for modest intramolecular affinity required to inhibit the PDZ2 domain binding site. Concurrently the transient interactions at the PDZ2 domain in the dynamic 'closed' state of NHERF1 (32) does not preclude rapid exchange between the myriad of binding partners during cell signaling.

One of major goals of this study was to probe the structural basis of impaired activity in naturally occurring NHERF1 mutations. What makes these disease mutations intriguing is the fact that none are physically located in the peptide-binding site and yet mediate interactions with binding partners. Using biophysical and structural tools we investigated in detail the complex relationship between the mutated PDZ domains, the ensemble dynamics of the intact scaffold and autoinhibition. By NMR, we were able to delineate the local structural changes triggered within the individual PDZ domains and evaluate their impact on the ensemble of global conformations in the intact scaffold by SANS. Complementary equilibrium and kinetic SPR measurements revealed the local changes in the PDZ domains influence the efficiency of substrate recognition by altering the rate of association/dissociation.

The disease mutations can detune NHERF1 function by at least two mechanisms. Either the altered structure of the mutant PDZ domains modulates the binding affinity to specific substrates (56), or the dynamic interaction between the PDZ2 and EB domain is impaired. Thus, one might expect mutations in the PDZ1 domain will primarily affect target

binding while those located in the PDZ2 domain have a dual impact on autoinhibition and binding affinity.

Therefore, it comes as no surprise that the binding affinity between ezFERM and the remote C-terminal EB domain shows marginal differences between the different mutations located in the PDZ domains and the wild-type protein. However, the kinetic measurements yield more interesting differences. The L110V mutant illustrates the effect of structural changes in the PDZ1 domain are not isolated and can mediate remotely the kinetics of EB domain docking in ezFERM. The two mutations, R153Q and E225K, in the PDZ2 domain present a very different conundrum. The intramolecular association even if perceptibly weakened in the mutants does not facilitate association with the EB domain. Instead, we observe the rate of association is much slower and that presents a potential kinetic bottleneck when anchoring NHERF1 at the membrane. A plausible explanation could be the EB domain is simply more disordered in the mutants when it is partially disengaged from PDZ2 (16). Thus, a higher entropic barrier must be overcome for binding induced folding of the C-terminal helical structure in the presence of ezFERM.

The magnitude of structural alterations in the mutated PDZ domains has a marginal impact on association between GRK6A peptide and is in sharp contrast to Npt2a where the interactions are debilitated (56). The two-step binding model with GRK6A, involves a high affinity PDZ1 binding step overlapped with much weaker association in the autoinhibited PDZ2 site. The structural rearrangement triggered by L110V and R153Q mutations results in a threefold decrease in the high affinity binding constant (nanomolar) for the GRK6A peptide but remains unchanged in E225K mutant. The weak binding step driven by the release of PDZ2 autoinhibition is uniformly depressed in the mutants but not eliminated. The complementary kinetic experiments from SPR revealed the rates of association and dissociation are significantly slower compared to the wild-type protein at both PDZ sites. The systematic retardation of the k_{on} rates in the mutants suggests an alternate mechanism to interfere with cellular activity.

From these key observations, we infer that the NHERF1 mutations could not only undermine interactions in some cases (56) but also deregulate time-dependent cellular events within signaling cascades as we have shown in this study. Notwithstanding the subtle differences in the intrinsic activity, all three NHERF1 mutants abolished PTH-sensitive P_i uptake resulting in comparable and severe phenotypes in patients (18). There is some evidence to suggest the pathogenic mechanism for heterozygous loss-of-function NHERF1 mutations identified in renal and skeletal phenotypes (56,57) and certain types of breast cancer (58) could be accounted for by haploinsufficiency (HI), wherein a single functional copy of a gene is insufficient to maintain normal function. Additional clinical and biochemical studies on other biological targets are necessary to generalize the conclusions from this study and the direction of future efforts. By combining different structural and biophysical techniques we present unique insight into molecular mechanisms by which the disease mutations can alter the function of a flexible scaffold like NHERF1 to assemble signaling complexes.

Experimental procedures

Protein Expression

The recombinant proteins corresponding to full length NHERF1 (residues 11-358) and the mutants (L110V, R153Q and E229K) were expressed in *E. coli* BL21 (DE3) cells (Novagen, Inc) and purified according to published methods (13). For uniform enrichment of the protein with $^{15}N/^{13}C$

isotopes, BL21 (DE3) cells were grown in M9-minimal medium containing $^{15}NH_4Cl$ (Cambridge Isotope Laboratories, CIL) and [$^{13}C_6$]-glucose as sole nitrogen and carbon source respectively. Deuterated proteins were produced by *E.coli* grown in M9-minimal medium containing 99% D_2O and protonated glucose. Typical NMR samples include 140 μM $^{2}H/^{13}C/^{15}N$ labeled NHERF1 dissolved in 20 mM Sodium Phosphate at pH 7.5, 150 mM NaCl, 0.5 mM dithiothreitol, 0.5 mM EDTA, 90% H_2O/10% D_2O. A 1:1 complex of $^{2}H/^{15}N/^{13}C$-labeled NHERF1 and unlabeled FERM domain from Ezrin was prepared under identical buffer conditions as described previously (13). The 22-residue C-terminal peptide from GRK6A with an N-terminal biotin tag was custom synthesized (Genscript, NJ).

NMR Spectroscopy

The NMR data was acquired on Bruker *AVANCE* series of spectrometers equipped with Z-axis gradient TCI/TXI CryoProbes^{TM} at a sample temperature of 15 °C and B_o field strengths of 800 and 900 MHz respectively. The multidimensional datasets were processed in Topspin 2.1 from Bruker Biospin and analyzed using CARA1.5 (59). The backbone resonances of $^{2}H/^{13}C/^{15}N$ labeled NHERF1 in the presence and absence of ezFERM domain and the three NHERF1 mutants were assigned by combining the information of the isolated fragments PDZ1 (BMRB 18824) + PDZ2CT (BMRB 16638) supported by a complete suite of TROSY based triple resonance experiments to assign the spacer between the PDZ domains. Two dimensional lineshape analysis was done in Topspin 3.5, by processing the data with gaussian window function (LB=-2Hz, GB=0.2) followed by deconvolution.

Small angle neutron scattering (SANS) experiments

SANS data were collected using the EQ-SANS instrument (60), which is a time-of-flight SANS instrument located at the Spallation Neutron Source (SNS) of Oak Ridge National Laboratory. A single instrument configuration with a 4m sample-to-detector distance was employed. The instrument choppers ran at 60 Hz and were set to provide a minimum wavelength of 2.5Å. The beam was defined with a 25mm diameter source aperture and a 10mm diameter sample aperture. The configuration spans a q-range from ~0.01 $Å^{-1}$ to ~0.40 $Å^{-1}$ (where $q=4\pi\sin(\theta)/\lambda$, 2θ the scattering angle, and λ the wavelength of the neutron), which probes the length scales required to see the complex and facilitates subtraction of the solvent background from the scattering signal from the samples. The SANS experiments were performed at 10 °C. The software Mantid was used to reduce the data from the samples and the associated solvent backgrounds using standard procedures that correct for incident flux spectrum, sample transmission and detector sensitivity, as well as the detector dark current (61). The final I(q) vs. q plots were obtained by azimuthally averaging the data after subtracting the solvent scattering. The reduced scattering data was subjected to consecutive Guinier analysis at $qR_g<1.13$ (**Figure 6B**). The length distribution function P(r) in **Figure 6C** was generated using the program GNOM (62).

Ensemble Analysis of SANS data

The initial structure of full-length NHERF1 'closed' structure was built in MDA (Multi Domain Alignment) Chimera 1.11 (63) by joining the PDZ1 and PDZ2CT template structures we determined previously by solution NMR (16,17). A representative ensemble of the two major states was generated in Xplor-NIH 2.47 by adapting established protocols for SAXS refinement of multidomain proteins detailed elsewhere (64) (65). Briefly PDZ1 and PDZ2 domains are treated as rigid bodies which can translate and rotate freely within a sphere of radius 40-80 Å during the Monte Carlo (MC) search

('randomizeDomainPos') by breaking the linkers. This step is followed by rejoining the linkers and a quick round of simulated annealing with torsion angle dynamics (TAD) (1000K to 25K) in the linkers followed by energy minimization. For the 'closed' structure the coordinates of PDZ2 and EB domains are grouped and treated as a single rigid body while in the 'open' state the EB domain can move freely in space with respect to the two PDZ domains. A van der Waal's repulsion term is used to prevent atomic clashes during the MC search and structures that violated the term rejected during the Powell energy minimization step. To prevent overlap of backbone atoms, the final ensemble of ~20K structures was generated in multiple steps by randomizing the domain positions of seed structures selected from the previous iteration. All the calculations were performed using either an inhouse multicore cluster or the Virtual Cluster NMRBox at UConn Health. The SANS scattering intensity curves were calculated for each conformer in the two ensemble pools corresponding to the 'open' and 'closed' states in CRYSON (66). Using a genetic clustering algorithm GAJOE from EOM 2.0 (36), the weighted average of the SANS curves for a subset of conformers from the randomized ensemble was fitted to the experimental data by minimizing the chi-square. Both the ensemble size (maximum 20 conformers) and the pool fractions ('open' and 'closed' states) were allowed to vary during the optimization process (100 rounds of genetic clustering). The program uses the optimized ensemble (to extract the average flexibility of the selected conformers, radius of gyration (R_g) and maximum length (D_{max}) distributions displayed in **Figure 6**. The overlay of the experimental and theoretical scattering intensity of the ensemble is shown in **Figure S5**.

Surface plasmon resonance experiments

Surface plasmon resonance (SPR) experiments were performed on an X100 instrument (GE Healthcare Life Sciences). To immobilize GRK6Act22, the Biacore CM5 Biosensor chips was activated by N-hydroxysuccinimide and N-ethyl-N'-[3-(diethylamino) propyl] carbodimide (GE Healthcare Life Sciences). The ligand, GRK6Act22, dissolved at 5 µg/ml in 10mM sodium acetate pH=4.9, was injected to coat the activated sensor chip surface. Free ligand was washed away, and uncoated sites blocked by 1 M ethanolamine, pH=8.5. The control flow cell was activated and blocked without ligand injection. The biotinylated GRK6Act22 ligand was immobilized on SA streptavidin sensor chip (GE Healthcare Life Sciences) using the manufacturer's protocol until the desired response target was reached. The analytes, NHERF1(wt) or mutants, were dissolved in HBS-EP buffer containing 10 mM HEPES buffer, pH=7.4, 3mM EDTA and 0.005% surfactant polysorbate 20 and different NaCl concentrations, respectively. The analytes were injected over the ezFERM coated surfaces at 30 µl/min for 180 seconds at increasing concentrations. The dissociation time is 800 seconds. At the end of injection-dissociation cycle, the sensor chip was regenerated with 4.0 M $MgCl_2$, 50 mM triethylamine (pH=9.15), and HBS-EP buffer. The SPR experiments were performed at 15 °C. The response curves were obtained by subtracting the background signal, generated from a control cell injected with the same analyte but without ligand coating of the hydrogel matrix to remove the effects of bulk refractive index and nonspecific binding. The data were further corrected by subtracting the signal generated from buffer alone. Background-corrected response curves were fit to the nonlinear equation one-site or two-site binding model in the BIA Evaluation Software. To obtain the kinetic association and dissociation rate constants, the SPR response curves (**Figure S6 and Figure S7**) were fitted using kinetic binding models provided by the BIA Evaluation Software.

Acknowledgement

This research was funded by NSF MCB-1817684 to ZB, 1R01DK105811 to PAF, and NIH and 2G12 RR003060 from the National Center for Research Resources to CCNY. The data collected at NYSBC was made possible by a grant from ORIP/NIH facility improvement grant CO6RR015495 and S10OD016432 (800 MHz). The 900 MHz NMR spectrometers were purchased with funds from NIH grant P41GM066354, the Keck Foundation, New York State Assembly, and U.S. Dept. of Defense.

This study made use of NMRbox: National Center for Biomolecular NMR Data Processing and Analysis, a Biomedical Technology Research Resource (BTRR), which is supported by NIH grant P41GM111135 (NIGMS).
A portion of the research conducted at ORNL's Spallation Neutron Source was sponsored by the Scientific User Facilities Division, Office of Basic Energy Sciences, U.S. Department of Energy. The authors thank Carrie Gao for technical support during the SANS experiments.

Conflict of interest: The authors declare that they have no conflicts of interest with the contents of this article

References

1. Garbett, D., and Bretscher, A. (2014) The surprising dynamics of scaffolding proteins. *Mol Biol Cell* **25**, 2315-2319.
2. Pawson, T., and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains. *Science* **300**, 445-452.
3. Harris, B. Z., and Lim, W. A. (2001) Mechanisms and role of PDZ domains in signalling complex assembly. *J. Cell. Sci.* **114**, 3219-3231.
4. Bhattacharyya, R. P., Remenyi, A., Yeh, B. J., and Lim, W. A. (2006) Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu Rev Biochem* **75**, 655-680.
5. Fanning, A. S., and Anderson, J. M. (1999) Protein modules as organizers of membrane structure. *Curr Opin Cell Biol* **11**, 432-439.
6. Morales, F. C., Takahashi, Y., Momin, S., Adams, H., Chen, X., and Georgescu, M. M. (2007) NHERF1/EBP50 head-to-tail intramolecular interaction masks association with PDZ domain ligands. *Mol Cell Biol* **27**, 2527-2537.
7. Li, J., Callaway, D. J., and Bu, Z. (2009) Ezrin induces long-range interdomain allostery in the scaffolding protein NHERF1. *J. Mol. Biol.* **392**, 166-180.
8. Farago, B., Li, J., Cornilescu, G., Callaway, D. J., and Bu, Z. (2010) Activation of nanoscale allosteric protein domain motion revealed by neutron spin echo spectroscopy. *Biophysical Journal* **99**, 3473-3482.
9. Haggie, P. M., Kim, J. K., Lukacs, G. L., and Verkman, A. S. (2006) Tracking of quantum dot-labeled CFTR shows near immobilization by C-terminal PDZ interactions. *Mol Biol Cell* **17**, 4937-4945.
10. Ardura, J. A., and Friedman, P. A. (2011) Regulation of G protein-coupled receptor function by Na+/H+ exchange regulatory factors. *Pharmacol Rev* **63**, 882-900.
11. Loureiro, C. A., Matos, A. M., Dias-Alves, A., Pereira, J. F., Uliyakina, I., Barros, P., Amaral, M. D., and Matos, P. (2015) A molecular switch in the scaffold NHERF1 enables misfolded CFTR to evade the peripheral quality control checkpoint. *Sci Signal* **8**, ra48.
12. Callaway, D. J., and Bu, Z. (2017) Visualizing the nanoscale: protein internal dynamics and neutron spin echo spectroscopy. *Curr Opin Struct Biol* **42**, 1-5.
13. Li, J., Dai, Z., Jana, D., Callaway, D. J., and Bu, Z. (2005) Ezrin controls the macromolecular complexes formed between an adapter protein Na+/H+ exchanger regulatory factor and the cystic fibrosis transmembrane conductance regulator. *J Biol Chem* **280**, 37634-37643.
14. Deshmukh, L., Schwieters, C. D., Grishaev, A., Ghirlando, R., Baber, J. L., and Clore, G. M. (2013) Structure and dynamics of full-length HIV-1 capsid protein in solution. *J Am Chem Soc* **135**, 16133-16147.
15. Bhattacharya, S., Lee, Y. T., Michowski, W., Jastrzebska, B., Filipek, A., Kuznicki, J., and Chazin, W. J. (2005) The modular structure of SIP facilitates its role in stabilizing multiprotein assemblies. *Biochemistry* **44**, 9462-9471.
16. Bhattacharya, S., Dai, Z., Li, J., Baxter, S., Callaway, D. J., Cowburn, D., and Bu, Z. (2010) A conformational switch in the scaffolding protein NHERF1 controls autoinhibition and complex formation. *J Biol Chem.* **285**, 9981-9994.

17. Bhattacharya, S., Ju, J. H., Orlova, N., Khajeh, J. A., Cowburn, D., and Bu, Z. (2013) Ligand-induced dynamic changes in extended PDZ domains from NHERF1. *J Mol Biol* **425**, 2509-2528.
18. Karim, Z., Gerard, B., Bakouh, N., Alili, R., Leroy, C., Beck, L., Silve, C., Planelles, G., Urena-Torres, P., Grandchamp, B., Friedlander, G., and Prie, D. (2008) NHERF1 Mutations and Responsiveness of Renal Parathyroid Hormone. *N Engl J Med* **359**, 1128-1135.
19. Georgescu, M. M., Morales, F. C., Molina, J. R., and Hayashi, Y. (2008) Roles of NHERF1/EBP50 in cancer. *Curr Mol Med* **8**, 459-468.
20. Lau, A. G., and Hall, R. A. (2001) Oligomerization of NHERF-1 and NHERF-2 PDZ domains: differential regulation by association with receptor carboxyl-termini and by phosphorylation. *Biochemistry* **40**, 8572-8580.
21. Li, J., Poulikakos, P. I., Dai, Z., Testa, J. R., Callaway, D. J., and Bu, Z. (2007) Protein kinase C phosphorylation disrupts Na+/H+ exchanger regulatory factor 1 autoinhibition and promotes cystic fibrosis transmembrane conductance regulator macromolecular assembly. *J Biol Chem* **282**, 27086-27099.
22. Garbett, D., LaLonde, D. P., and Bretscher, A. (2010) The scaffolding protein EBP50 regulates microvillar assembly in a phosphorylation-dependent manner. *J Cell Biol* **191**, 397-413.
23. Hall, R. A., Spurney, R. F., Premont, R. T., Rahman, N., Blitzer, J. T., Pitcher, J. A., and Lefkowitz, R. J. (1999) G protein-coupled receptor kinase 6A phosphorylates the Na(+)/H(+) exchanger regulatory factor via a PDZ domain-mediated interaction. *J Biol Chem* **274**, 24328-24334.
24. Zhang, Q., Xiao, K., Paredes, J. M., Mamonova, T., Sneddon, W. B., Liu, H., Wang, D., Li, S., McGarvey, J. C., Uehling, D., Al-Awar, R., Joseph, B., Jean-Alphonse, F., Orte, A., and Friedman, P. A. (2019) Parathyroid hormone initiates dynamic NHERF1 phosphorylation cycling and conformational changes that regulate NPT2A-dependent phosphate transport. *J Biol Chem* **294**, 4546-4571.
25. Shen, Y., and Bax, A. (2015) Protein structural information derived from NMR chemical shift with the neural network program TALOS-N. *Methods Mol Biol* **1260**, 17-32
26. Nielsen, J. T., and Mulder, F. A. A. (2018) POTENCI: prediction of temperature, neighbor and pH-corrected chemical shifts for intrinsically disordered proteins. *J Biomol NMR* **70**, 141-165.
27. Mamonova, T., Kurnikova, M., and Friedman, P. A. (2012) Structural basis for NHERF1 PDZ domain binding. *Biochemistry* **51**, 3110-3120.
28. Garbett, D., and Bretscher, A. (2012) PDZ interactions regulate rapid turnover of the scaffolding protein EBP50 in microvilli. *The Journal of Cell Biology* **198**, 195-203.
29. Laederach, A., Cradic, K. W., Fulton, D. B., and Andreotti, A. H. (2003) Determinants of intra versus intermolecular self-association within the regulatory domains of Rlk and Itk. *J Mol Biol* **329**, 1011-1020.
30. Borcherds, W., Becker, A., Chen, L., Chen, J., Chemes, L. B., and Daughdrill, G. W. (2017) Optimal Affinity Enhancement by a Conserved Flexible Linker Controls p53 Mimicry in MdmX. *Biophys J* **112**, 2038-2042.
31. Zhou, H. X. (2006) Quantitative relation between intermolecular and intramolecular binding of pro-rich peptides to SH3 domains. *Biophys J* **91**, 3170-3181.

32. Cheng, H., Li, J., Fazlieva, R., Dai, Z., Bu, Z., and Roder, H. (2009) Autoinhibitory interactions between the PDZ2 and C-terminal domains in the scaffolding protein NHERF1. *Structure* **17**, 660-669.
33. Jayasundar, J. J., Ju, J. H., He, L., Liu, D., Meilleur, F., Zhao, J., Callaway, D. J., and Bu, Z. (2012) Open conformation of ezrin bound to phosphatidylinositol 4,5-bisphosphate and to F-actin revealed by neutron scattering. *J Biol Chem* **287**, 37119-37133.
34. Reczek, D., and Bretscher, A. (1998) The carboxyl-terminal region of EBP50 binds to a site in the amino-terminal domain of ezrin that is masked in the dormant molecule. *J Biol Chem* **273**, 18452-18458.
35. Holcomb, J., Jiang, Y., Guan, X., Trescott, L., Lu, G., Hou, Y., Wang, S., Brunzelle, J., Sirinupong, N., Li, C., and Yang, Z. (2014) Crystal structure of the NHERF1 PDZ2 domain in complex with the chemokine receptor CXCR2 reveals probable modes of PDZ2 dimerization. *Biochem Biophys Res Commun* **448**, 169-174.
36. Tria, G., Mertens, H. D., Kachala, M., and Svergun, D. I. (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207-217.
37. Bernado, P., Mylonas, E., Petoukhov, M. V., Blackledge, M., and Svergun, D. I. (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. *J Am Chem Soc* **129**, 5656-5664.
38. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A., and Sali, A. (2016) FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. *Nucleic Acids Res* **44**, W424-429.
39. Perkins, S. J., Wright, D. W., Zhang, H., Brookes, E. H., Chen, J., Irving, T. C., Krueger, S., Barlow, D. J., Edler, K. J., Scott, D. J., Terrill, N. J., King, S. M., Butler, P. D., and Curtis, J. E. (2016) Atomistic modelling of scattering data in the Collaborative Computational Project for Small Angle Scattering (CCP-SAS). *J Appl Crystallogr* **49**, 1861-1875.
40. Callaway, D. J. E., Matsui, T., Weiss, T., Stingaciu, L. R., Stanley, C. B., Heller, W. T., and Bu, Z. (2017) Controllable Activation of Nanoscale Dynamics in a Disordered Protein Alters Binding Kinetics. *J Mol Biol* **429**, 987-998.
41. Farago, B., Li, J., Cornilescu, G., Callaway, D. J., and Bu, Z. (2010) Activation of nanoscale allosteric protein domain motion revealed by neutron spin echo spectroscopy. *Biophys J* **99**, 3473-3482.
42. Mamonova, T., Zhang, Q., Chandra, M., Collins, B. M., Sarfo, E., Bu, Z., Xiao, K., Bisello, A., and Friedman, P. A. (2017) Origins of PDZ Binding Specificity. A Computational and Experimental Study Using NHERF1 and the Parathyroid Hormone Receptor. *Biochemistry* **56**, 2584-2593.
43. Finnerty, C. M., Chambers, D., Ingraffea, J., Faber, H. R., Karplus, P. A., and Bretscher, A. (2004) The EBP50-moesin interaction involves a binding site regulated by direct masking on the FERM domain. *J Cell Sci* **117**, 1547-1552.
44. Terawaki, S., Maesaki, R., and Hakoshima, T. (2006) Structural basis for NHERF recognition by ERM proteins. *Structure* **14**, 777-789.
45. Gianni, S., Engström, Å., Larsson, M., Calosci, N., Malatesta, F., Eklund, L., Ngang, C. C., Travaglini-Allocatelli, C., and Jemth, P. (2005) The Kinetics of PDZ Domain-Ligand Interactions and Implications for the Binding Mechanism. *Journal of Biological Chemistry* **280**, 34805-34812.

46. Jemth, P., and Gianni, S. (2007) PDZ domains: folding and binding. *Biochemistry* **46**, 8701-8708.
47. Feng, W., and Zhang, M. (2009) Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. *Nat Rev Neurosci* **10**, 87-99.
48. Long, J. F., Tochio, H., Wang, P., Fan, J. S., Sala, C., Niethammer, M., Sheng, M., and Zhang, M. (2003) Supramodular structure and synergistic target binding of the N-terminal tandem PDZ domains of PSD-95. *J. Mol. Biol.* **327**, 203-214.
49. Wang, W., Weng, J., Zhang, X., Liu, M., and Zhang, M. (2009) Creating conformational entropy by increasing interdomain mobility in ligand binding regulation: a revisit to N-terminal tandem PDZ domains of PSD-95. *J. AM. Chem. Soc.* **131**, 787-796.
50. George, R. A., and Heringa, J. (2002) An analysis of protein domain linkers: their classification and role in protein folding. *Protein Eng* **15**, 871-879.
51. Theillet, F. X., Kalmar, L., Tompa, P., Han, K. H., Selenko, P., Dunker, A. K., Daughdrill, G. W., and Uversky, V. N. (2013) The alphabet of intrinsic disorder: I. Act like a Pro: On the abundance and roles of proline residues in intrinsically disordered proteins. *Intrinsically Disord Proteins* **1**, e24360.
52. Uversky, V. N. (2013) The alphabet of intrinsic disorder: II. Various roles of glutamic acid in ordered and intrinsically disordered proteins. *Intrinsically Disord Proteins* **1**, e24684.
53. Uversky, V. N. (2015) The intrinsic disorder alphabet. III. Dual personality of serine. *Intrinsically Disord Proteins* **3**, e1027032.
54. Johnson, L. N., and Lewis, R. J. (2001) Structural basis for control by phosphorylation. *Chem Rev* **101**, 2209-2242.
55. Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z., and Dunker, A. K. (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res* **32**, 1037-1049.
56. Wang, B., Means, C. K., Yang, Y., Mamonova, T., Bisello, A., Altschuler, D. L., Scott, J. D., and Friedman, P. A. (2012) Ezrin-anchored protein kinase A coordinates phosphorylation-dependent disassembly of a NHERF1 ternary complex to regulate hormone-sensitive phosphate transport. *J Biol Chem.* **287**, 24148-24163.
57. Shenolikar, S., Voltz, J. W., Minkoff, C. M., Wade, J. B., and Weinman, E. J. (2002) Targeted disruption of the mouse NHERF-1 gene promotes internalization of proximal tubule sodium-phosphate cotransporter type IIa and renal phosphate wasting. *Proc Natl Acad Sci U S A* **99**, 11470-11475.
58. Pan, Y., Wang, L., and Dai, J. L. (2006) Suppression of breast cancer cell growth by Na+/H+ exchanger regulatory factor 1 (NHERF1). *Breast Cancer Res* **8**, R63.
59. Keller, R. (2004) *Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment.* PhD, ETH.
60. Zhao, J. K., Gao, C. Y., and Liu, D. (2010) The extended Q-range small-angle neutron scattering diffractometer at the SNS. *Journal of Applied Crystallography* **43**, 1068-1077.
61. Arnold, O., Bilheux, J.-C., Borreguero, J., Buts, A., Campbell, S. I., Chapon, L., Doucet, M., Draper, N., Leal, R. F., and Gigg, M. (2014) Mantid—data analysis and visualization package for neutron scattering and μ SR experiments. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **764**, 156-166.

62. Semenyuk, A. V., and Svergun, D. I. (1991) GNOM - A program Package for small-angle scattering data-processing. *J. Appl. Cryst.* **24**, 537-540.
63. Hertig, S., Goddard, T. D., Johnson, G. T., and Ferrin, T. E. (2015) Multidomain Assembler (MDA) Generates Models of Large Multidomain Proteins. *Biophys J* **108**, 2097-2102.
64. Deshmukh, L., Schwieters, C. D., Grishaev, A., and Clore, G. M. (2016) Quantitative Characterization of Configurational Space Sampled by HIV-1 Nucleocapsid Using Solution NMR, X-ray Scattering and Protein Engineering. *Chemphyschem* **17**, 1548-1552.
65. Schwieters, C. D., Bermejo, G. A., and Clore, G. M. (2018) Xplor-NIH for molecular structure determination from NMR and other data sources. *Protein Sci* **27**, 26-40.
66. Svergun, D. I., Richard, S., Koch, M. H., Sayers, Z., Kuprin, S., and Zaccai, G. (1998) Protein hydration in solution: experimental observation by x-ray and neutron scattering. *Proc Natl Acad Sci U S A* **95**, 2267-2272.

Abbreviations: CB loop, carboxylate binding loop; CFTR, cystic fibrosis transmembrane conductance regulator; EB, ezrin binding; ERM, ezrin-radixin-moesin; FERM, 4.1-ezrin/radixin/moesin; CT, C-terminal; NHERF, Na^+/H^+ exchange regulatory factor; NMR, Nuclear Magnetic Resonance; NOE, nuclear Overhauser effect; NPT2A, Sodium-dependent phosphate transport protein 2A; SAXS, small-angle X-ray scattering; SANS, small-angle neutron scattering; SD, Standard Deviation; SPR, surface plasmon resonance; TROSY, Transverse relaxation optimized spectroscopy; PTH, parathyroid hormone.

Table 1. Summary of ensemble averaged structural and dynamic parameters from SANS data analysis of NHERF1 in GAJOE.

	C(mg/ml)	I(0)(1/cm)	R_g (Å)	D_{max} (Å)	$R_{flex,selected}$ (%)	R_{sigma} (%)	χ^2
NHERF1(wt)[a]			40.4	135			
NHERF1(wt)[b]	1.0	0.1565 ± 0.0040	39.2	134.4	83	0.9	0.68
L110V[b]	0.95	0.1437 ± 0.0043	39.7	134.8	78	0.76	0.75
R153Q[b]	1.13	0.1682 ± 0.0044	39.5	124.4	69	0.57	0.73
E225K[b]	1.11	0.1674 ± 0.0041	35.5	121.2	72	0.65	0.76

(a) R_g/D_{max} from SEC-SAXS data reported previously (Callaway et al., 2017). (b) R_g/D_{max} and flexibility parameters from final ensemble obtained by fitting SANS data in GAJOE. Starting ensemble composed of random structures yields R_{flex} = 85.5%.

Figure Captions

Figure 1: A) The amino acid sequence (residues 1-358) of full-length Human NHERF1 protein annotated with the secondary structure in the extended PDZ1 (residues 11-112), PDZ2 (residues 150-260) domains, and the C-terminal Ezrin binding domain (residues 329-358). The color-coded secondary structure elements highlight the α-helices (green), β-strands (blue), COO^- binding-loop (red) and the C-terminal Type 1 PDZ binding motif (red). The two longer fragments of NHERF1 discussed in the text include PDZ2CT domains (residues 150-358) and CT domain (residues 260-358) respectively.

Figure 2. (A) Overlay of 2D N^{15}-TROSY spectra of full-length $^{2}H/^{13}C/^{15}N$-labeled NHERF1 bound to unlabeled Ezrin (black) and the free-state (red). (B) Bar plot of weighted amide proton ($^{1}H^{N}$) and nitrogen (^{15}N) chemical shift difference between the bound and free protein versus residue number. The red line shows the mean±SD for CSD (0.062 ppm). (C) Secondary structure prediction of NHERF1 based on chemical shift index calculated in TALOS+ (25). Residues in β-sheet (red) and α-helices (blue) are indicated by different colors. The greyed box indicates the exchanged broadened residues in the peptide binding site at β2-strand from PDZ1 and PDZ2 domain respectively and the C-terminal EB domain.

Figure 3. Correlation plots of intrinsically disordered protein chemical shifts predicted by POTENCI (26) and the corresponding experimental shifts for residues in the linker between PDZ1-PDZ2 domain (residues 113-146) and the C-terminal region (residues 252-323). The panels display (A) Amide nitrogen (N^{15}), (B) Alpha carbon ($^{13}C_{\alpha}$), (C) Beta carbon ($^{13}C_{\beta}$), and (D) Carbonyl carbon ($^{13}C'$) shifts respectively.

Figure 4. (A) Overlay of 2D N^{15}-TROSY spectra of NHERF1 in the free (black) and GRK6A peptide bound (1:7 ratio) state (red) at 900 MHz and 15 °C. (B) Boxed inset from 2D N^{15}-TROSY overlay of three different titration points with NHERF1 and GRK6A peptide ratios indicated in the legend. The selected region shows the effect of peptide binding on the N-terminal Gly23/Gly163 from the CB loop (**Figure 1**) which is not exchange broadened in the apo-state of the protein. (C) Weighted average of amide CSD of NHERF1 bound to GRK6A peptide (1:7 ratio) peptide at 15 °C. The dashed horizontal line at 0.8 ppm (mean±SD) indicates the threshold for significant chemical shift perturbation.

Figure 5: Chemical shift perturbation analysis of NHERF1 mutants compared to the wild-type protein. The histograms display the weighted average of amide ($^{1}H/^{15}N$) chemical shift difference as a function of residue number from data acquired at 900 MHz and 15 °C. The panels display the data for three different mutants of full-length NHERF1; (A) L110V (0.16 ppm), (D) R153Q (0.087 ppm) and (G) E225K (0.026 ppm). The red dashed line indicates the mean±SD with values greater than the threshold mapped on the NMR structure of the isolated (B) PDZ1 (2MOT), and (E)-(H) PDZ2 (2KJD) domain structures. (C),(F),(I) Overlay of 2D N^{15}-TROSY spectrum of ezFERM bound NHERF1(wt) (blue), free NHERF1(wt) (black) and variants (red), which are indicated in the panel legends, to show the changes in Gly23/Gly163 from CB loop in the two PDZ domains. The corresponding line-width (Hz) at half-height of the Gly crosspeaks are similarly color-coded along F2 ($^{1}H^{N}$) and F1 (N^{15}) dimensions respectively.

Figure 6: Ensemble analysis of SANS data of wild-type and mutant NHERF1 protein structures using GAJOE/EOM. **(**A**)** I(q) vs q experimental data. The scale of scattering intensity is arbitrary to display all mutants. **(**B**)** Guinier plot calculated from I(q) vs q experimental data. **(**C**)** P(r) functions that are normalized by forward scattering intensity I(0) of the different mutants. (D) Overlay of R_g, and, (E) D_{max} frequency distribution of starting ensemble (black) and best fitted ensemble pools of different NHERF1 proteins. The displayed structures represent the major fractions from the optimized ensemble selected by GAJOE. (F)-(G) NHERF1(wt), (H)-(I) L110V, (J) R153Q, (K) E225K. The legend in each panel includes R_g, D_{max}, distance between the center-of-mass of the PDZ1 and PDZ2 domains, fraction of the conformer selected from the two starting pools.

Figure 7. Affinity and kinetics of NHERF1(wt) and mutants binding to ezFERM**.** (A) The equilibrium curves of NHERF1(wt) and mutants binding to ezFERM fitted with a 1:1 binding model, see **Table S1** for K_d values. Displayed univariate scatter plots obtained from three repeat SPR measurements: (B)(D) two-state kinetic association rate constants (k_{a1}, k_{a2}), and, (C)(E) dissociation rate constants (k_{d1}, k_{d2}) constants. Different symbols with Y-error bars have been used for each measurement.

Figure 8. Affinity and kinetics of NHERF1(wt) and mutants binding to GRK6A peptide. (A) The equilibrium curves of NHERF1(wt) and mutants binding to GRK6A peptide from SPR are fitted by a two-site binding model, see **Table S2** for K_d values. Displayed univariate scatter plots obtained from three repeat SPR measurements: (B)(D) bivalent association rate constants (k_{a1}, k_{a2}), and, (C)(E) dissociation rate constants (k_{d1}, k_{d2}). Different symbols with Y-error bars have been used for each measurement.

Figure 1

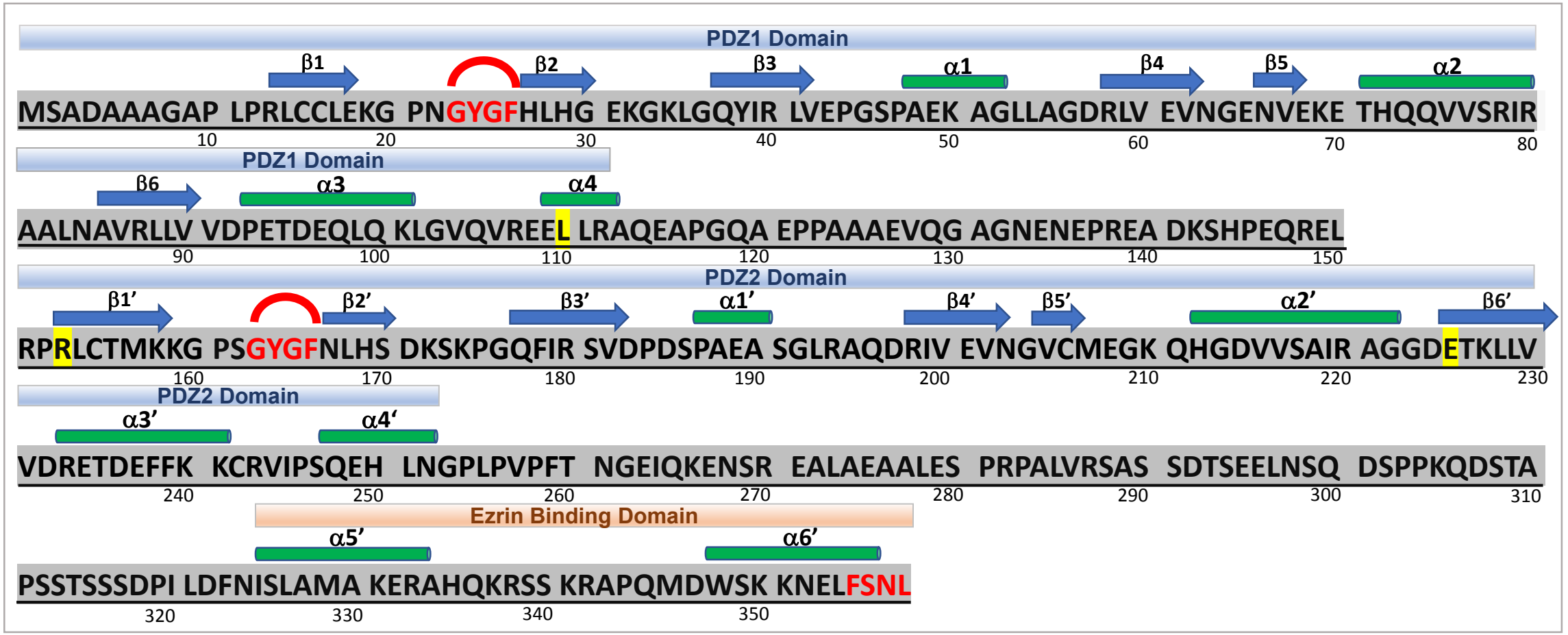

PDZ1 Domain
β1
β2
β3
α1
β4
β5
α2
MSADAAAGAP LPRLCCLEKG PNGYGFHLHG EKGKLGQYIR LVEPGSPAEK AGLLAGDRLV EVNGENVEKE THQQVVSRIR
10
20
30
40
50
60
70
80
PDZ1 Domain
β6
α3
α4
AALNAVRLLV VDPETDEQLQ KLGVQVREEL LRAQEAPGQA EPPAAAEVQG AGNENEPREA DKSHPEQREL
90
100
110
120
130
140
150
PDZ2 Domain
β1'
β2'
β3'
α1'
β4'
β5'
α2'
β6'
RPRLCTMKKG PSGYGFNLHS DKSKPGQFIR SVDPDSPAEA SGLRAQDRIV EVNGVCMEGK QHGDVVSAIR AGGDETKLLV
160
170
180
190
200
210
220
230
PDZ2 Domain
α3'
α4'
VDRETDEFFK KCRVIPSQEH LNGPLPVPFT NGEIQKENSR EALAEAALES PRPALVRSAS SDTSEELNSQ DSPPKQDSTA
240
250
260
270
280
290
300
310
Ezrin Binding Domain
α5'
α6'
PSSTSSSDPI LDFNISLAMA KERAHQKRSS KRAPQMDWSK KNELFSNL
320
330
340
350

Figure 2

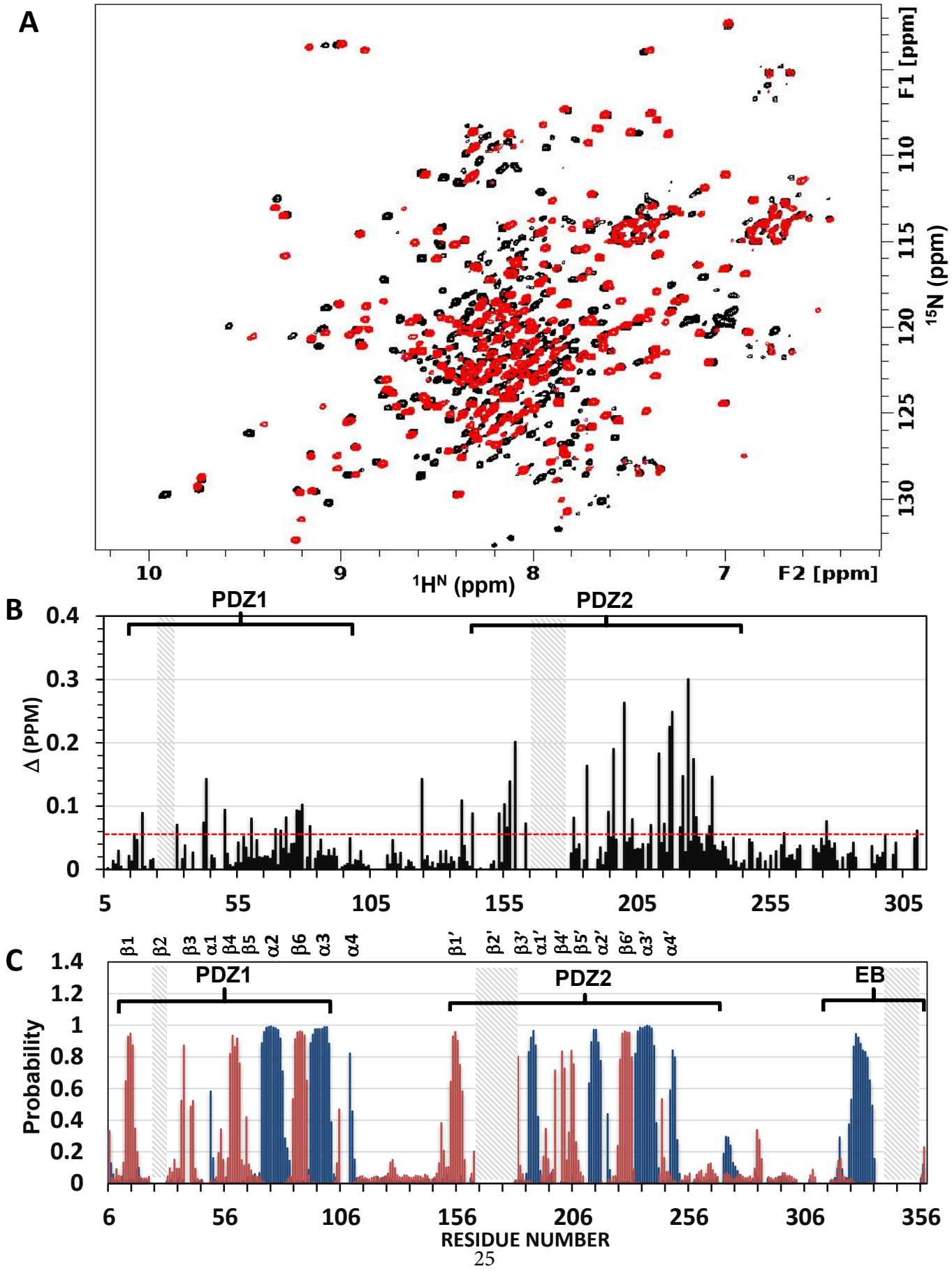

A
F1 [ppm]
110
115
120
125
130
15N (ppm)
10
9
8
7
F2 [ppm]
1HN (ppm)
B
PDZ1
PDZ2
0.4
0.3
0.2
0.1
0
Δ (PPM)
5
55
105
155
205
255
305
C
β1 β2 β3 α1 β4 β5 α2 β6 α3 α4 β1' β2' β3' α1' β4' β5' α2' β6' α3' α4'
PDZ1
PDZ2
EB
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Probability
6
56
106
156
206
256
306
356
RESIDUE NUMBER

Figure 3

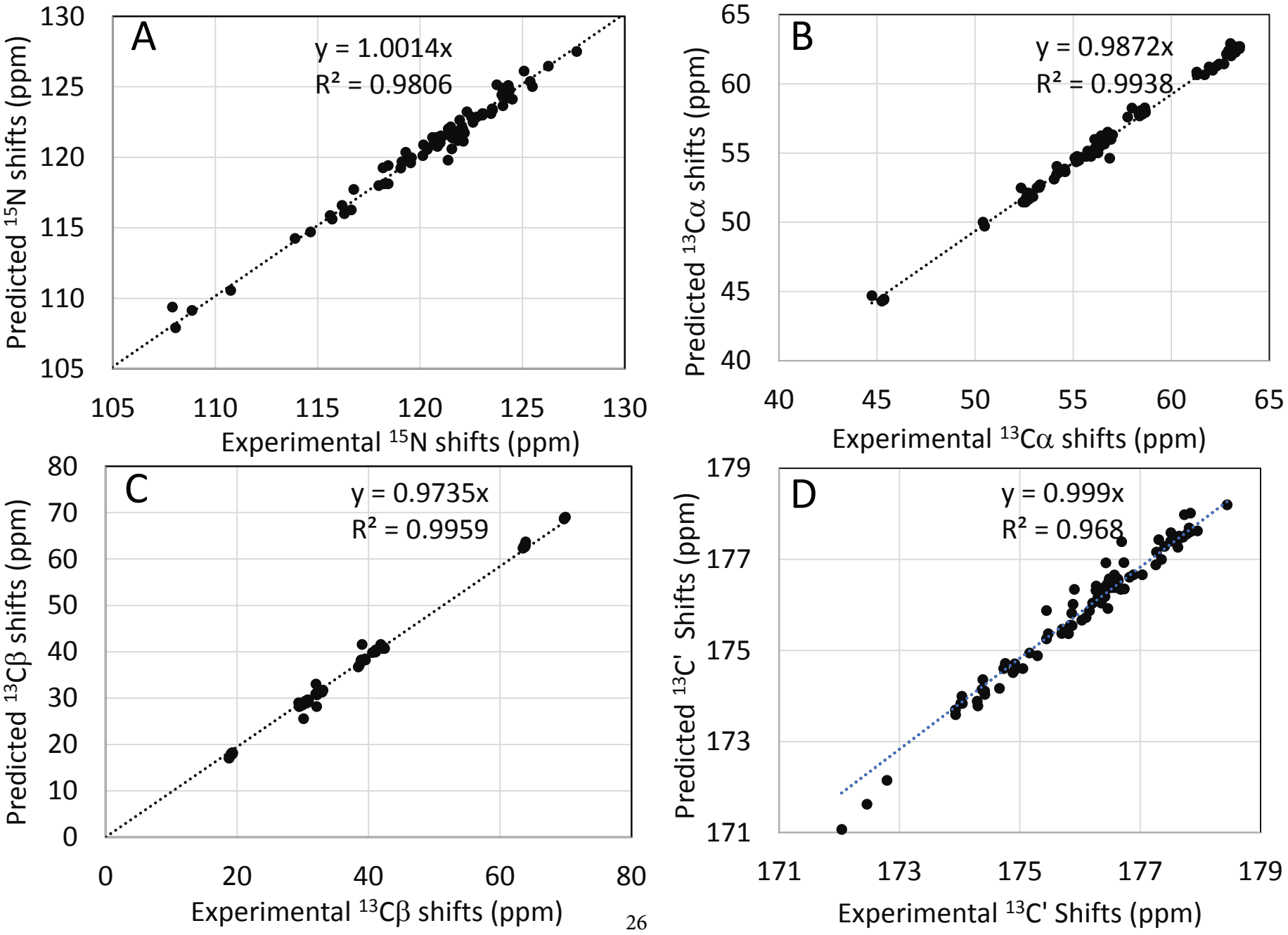

A
y = 1.0014x
R² = 0.9806
Predicted ¹⁵N shifts (ppm)
Experimental ¹⁵N shifts (ppm)
105
110
115
120
125
130
B
y = 0.9872x
R² = 0.9938
Predicted ¹³Cα shifts (ppm)
Experimental ¹³Cα shifts (ppm)
40
45
50
55
60
65
C
y = 0.9735x
R² = 0.9959
Predicted ¹³Cβ shifts (ppm)
Experimental ¹³Cβ shifts (ppm)
0
10
20
30
40
50
60
70
80
D
y = 0.999x
R² = 0.968
Predicted ¹³C' Shifts (ppm)
Experimental ¹³C' Shifts (ppm)
171
173
175
177
179

Figure 4

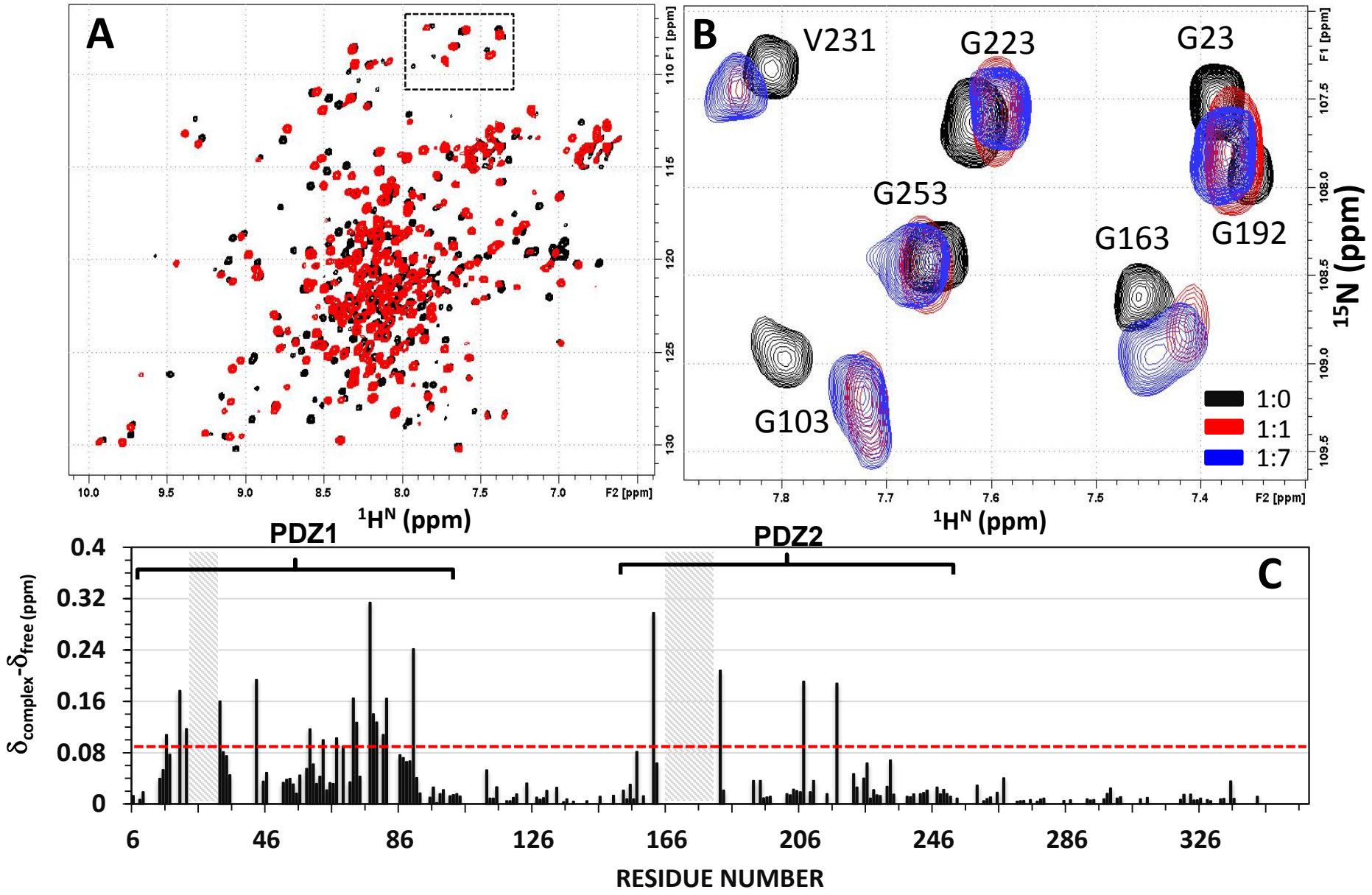

A
B
V231
G223
G23
G253
G163
G192
G103
1:0
1:1
1:7
^{15}N (ppm)
^{1}H^{N} (ppm)
^{1}H^{N} (ppm)
F2 [ppm]
F1 [ppm]
PDZ1
PDZ2
C
$\delta_{complex}$-δ_{free} (ppm)
RESIDUE NUMBER

Figure 5

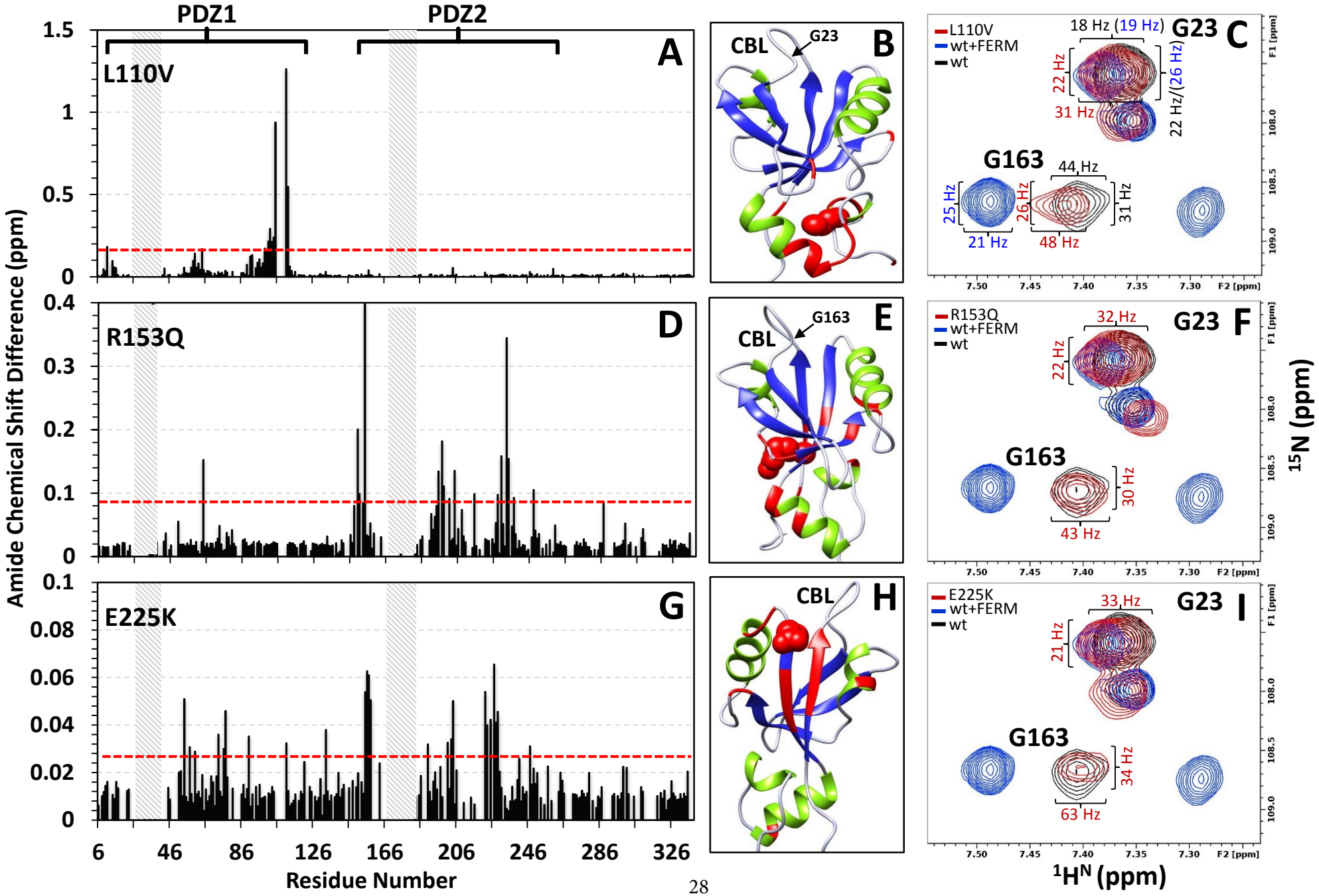
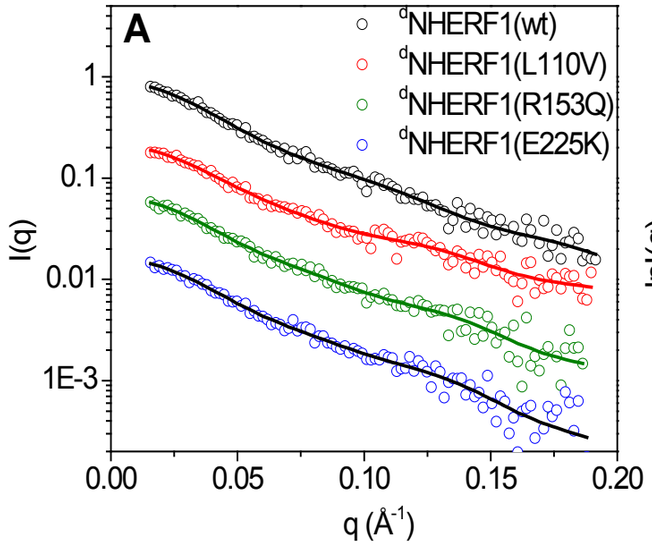
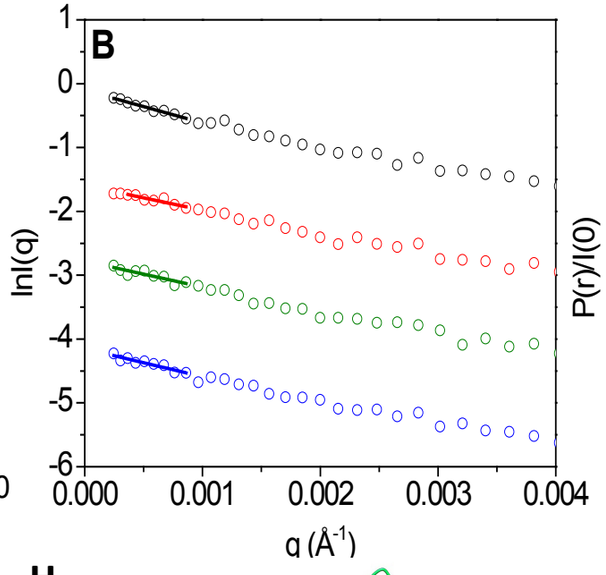
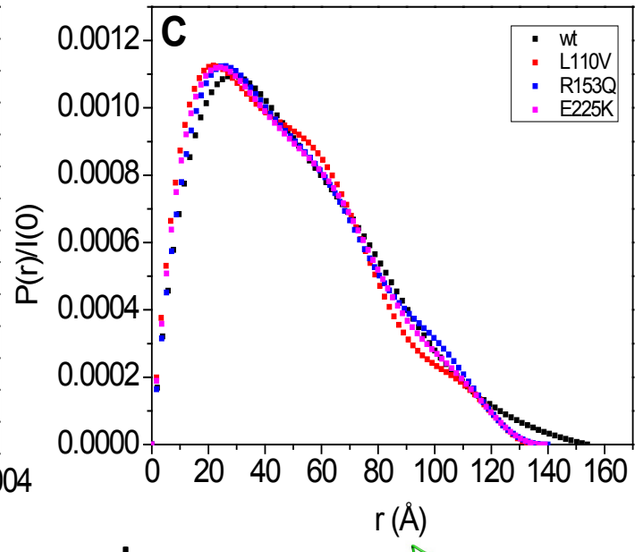
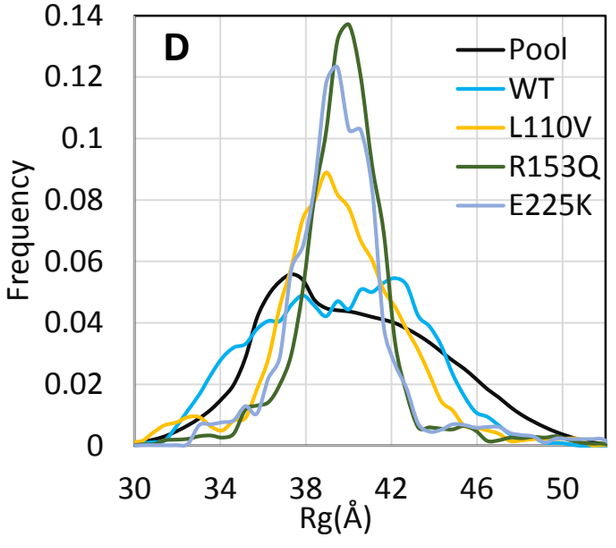

PDZ1
PDZ2
L110V
A
R153Q
D
E225K
G
Amide Chemical Shift Difference (ppm)
Residue Number
CBL
G23
B
G163
E
H
L110V
wt+FERM
wt
18 Hz (19 Hz)
22 Hz
31 Hz
22 Hz/(26 Hz)
G23
C
G163
44 Hz
25 Hz
26 Hz
31 Hz
21 Hz
48 Hz
R153Q
32 Hz
F
30 Hz
43 Hz
E225K
33 Hz
I
34 Hz
63 Hz
F1 [ppm]
F2 [ppm]
15N (ppm)
1HN (ppm)

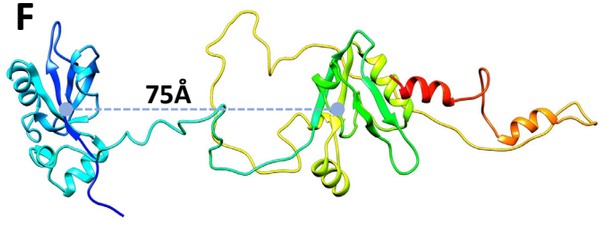
Figure 6



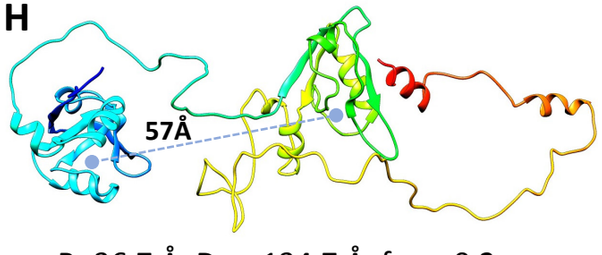




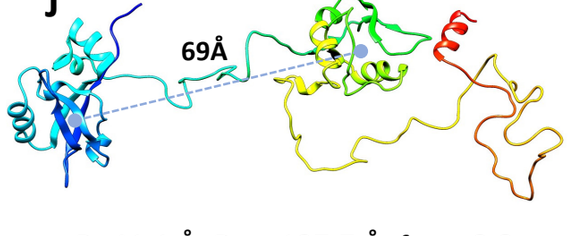




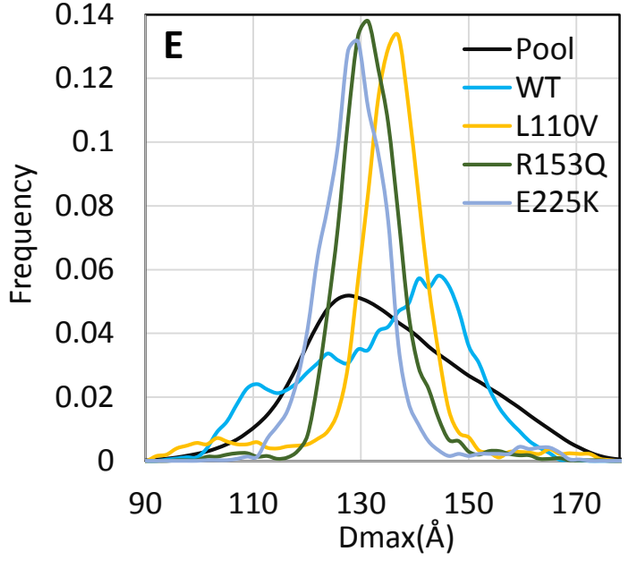
R_g 41.9 Å, D_{max} 148.3 Å, f_{close}=0.45

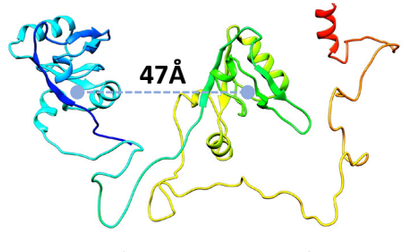


R_g 36.7 Å, D_{max} 134.7 Å, f_{close}=0.2

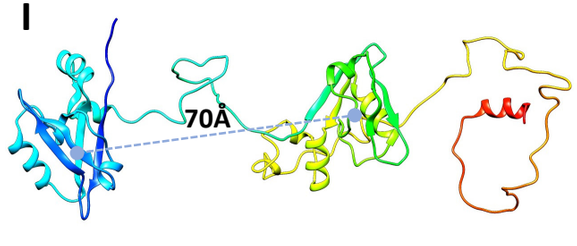


R_g 41.1 Å, D_{max} 127.7 Å, f_{open}=0.8

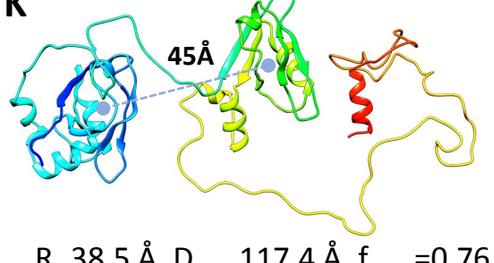




R_g 33.5 Å, D_{max} 111.0 Å, f_{open}=0.18

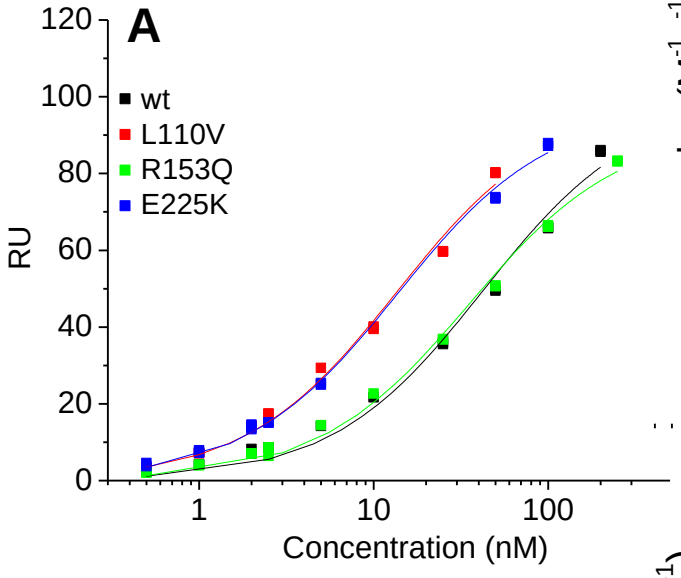


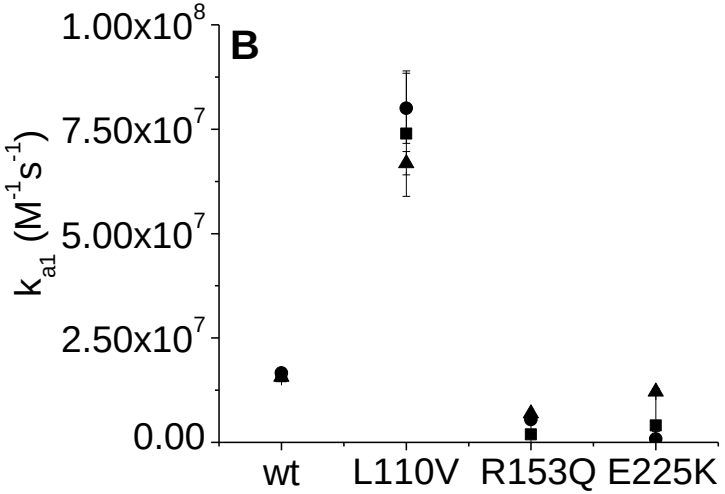
R_g 42.5 Å, D_{max} 136.9 Å, f_{open}=0.5

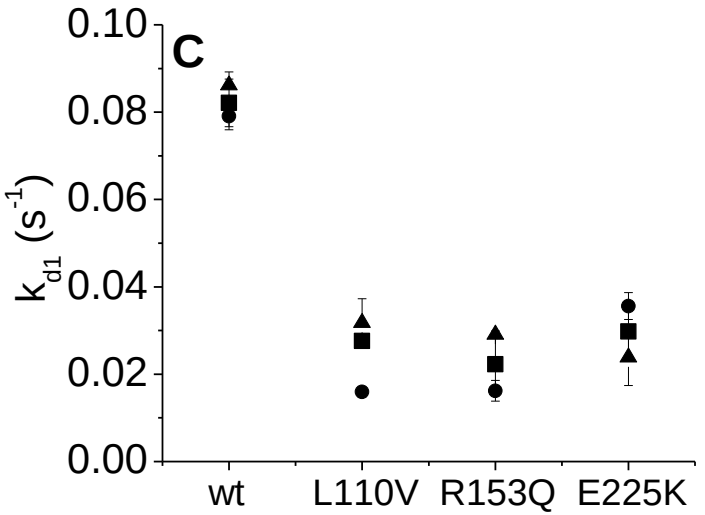


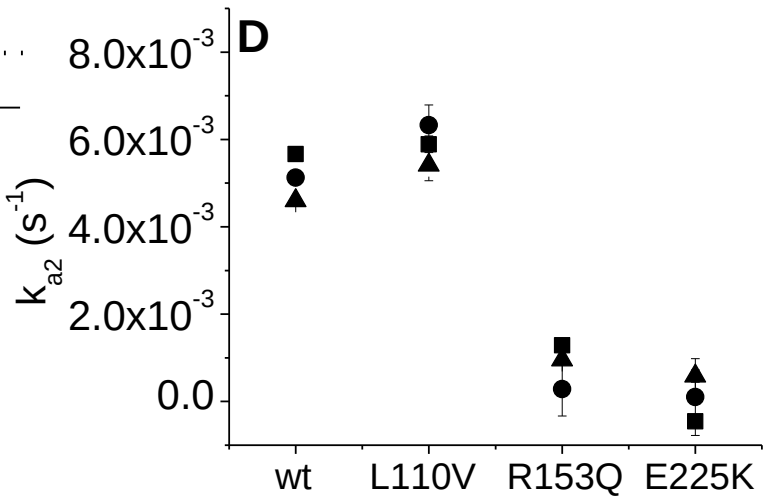
R_g 38.5 Å, D_{max} 117.4 Å, f_{open}=0.76

Figure 7


A
wt
L110V
R153Q
E225K
RU
Concentration (nM)


B
k_{a1} ($M^{-1}s^{-1}$)
wt
L110V
R153Q
E225K


C
k_{d1} (s^{-1})
wt
L110V
R153Q
E225K


D
k_{a2} (s^{-1})
wt
L110V
R153Q
E225K

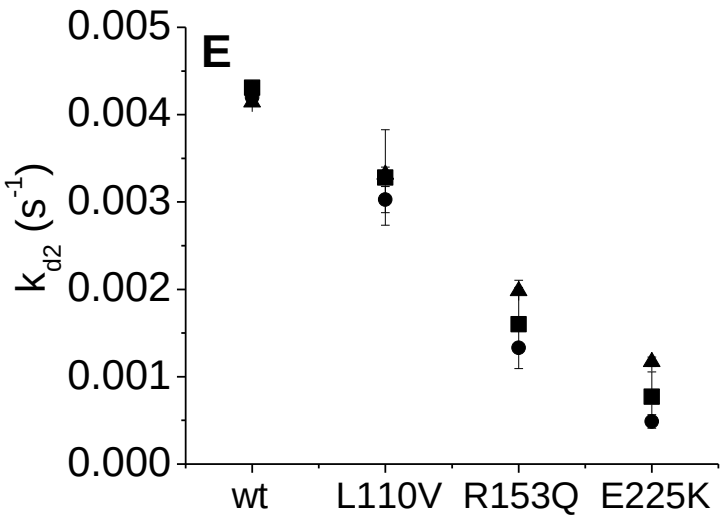
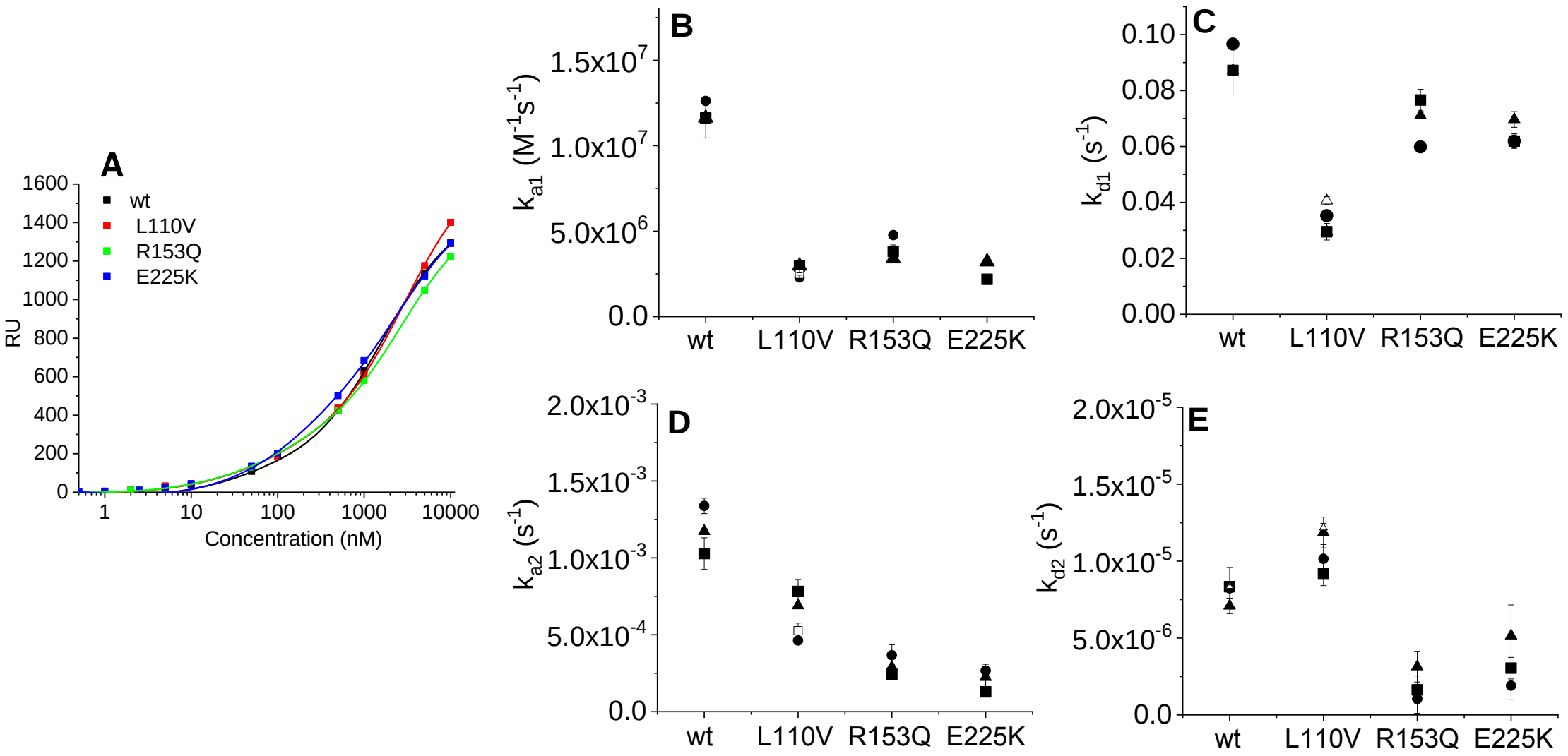

E
k_{d2} (s^{-1})
wt
L110V
R153Q
E225K

Figure 8


A
wt
L110V
R153Q
E225K
RU
Concentration (nM)
B
k_{a1} ($M^{-1}s^{-1}$)
C
k_{d1} (s^{-1})
D
k_{a2} (s^{-1})
E
k_{d2} (s^{-1})
wt
L110V
R153Q
E225K